Informational Frictions and the Life-Cycle Dynamics of Labor Market Outcomes

Georg Duernecker*

PRELIMINARY AND INCOMPLETE

Abstract

This paper studies the life-cycle dynamics of individual labor market outcomes. After entry into the labor market, young individuals typically change jobs very frequently and retain new jobs just for a short period of time. In later stages of their career, workers tend to hold stable jobs and they are substantially more likely to keep a new job than in the early years. This paper argues that the labor market experience individuals accumulate early in their career affects their job mobility in later stages. We construct a life-cycle model of the labor market whose main characteristic is an information imperfection in the matching process. The key ingredient is that the imperfection is assumed to be worker-specific and in particular it is linked to an individual's previous labor market history. We estimate the model by indirect inference on data from the *NLSY* 79 and find that it can capture very well the observed life-cycle profile of individual labor market mobility.

JEL classification: C15, D83, E24, J62

Keywords: Job mobility, Job turnover, Life-cycle, Signaling, Information imperfections, Indirect Inference

*Department of Economics, University of Mannheim, L7 3-5, 68161 Mannheim, Germany, email: duernecker@uni-mannheim.de

1 Introduction

Why does individual job mobility systematically differ across age groups? What we observe is that young individuals typically change jobs very frequently and retain each new job just for a short period of time. By contrast, experienced workers tend to hold stable jobs, but more importantly, they are also less likely to separate from a new job than the young. On the aggregate level, this leads to high rates of job turnover and unemployment among the young but comparatively moderate figures for prime-age workers. This pattern is a common characteristic of all labor markets in OECD countries, though it is particularly pronounced in the U.S.

The literature considers the initially high, but declining, turnover for the young to be the outcome of a search process in which workers experiment with jobs in order to find the right match. Accordingly, individuals hold a number of jobs in the first years after labor market entry. Eventually, after a series of failures, one job turns out to be a good match and lasts several years. This explanation is incomplete as it implies that each prime-age worker, who separates from a long-term job and seeks to find a new occupation would, necessarily, have to go through the same wasteful search process as at the beginning of the career. This is clearly not the case. Empirical research has uncovered that the job mobility of *newly employed* workers is declining with age since prime-age workers are significantly less likely to separate from a *new job* than young adults. Evidently, mismatch is a phenomenon that is strongly concentrated among young, but appears to be less common among job seekers of higher age cohorts.

The objective of this paper is to contribute to the understanding of the observed pattern of individual job mobility along the life-cycle. To this end, we first present new empirical evidence documenting the life-cycle dynamics of job mobility in the U.S. labor market. Second, in order to account for the observed pattern, we construct a life-cycle model of the labor market. The framework we propose formalizes the hypotheses that the labor market experience individuals accumulate in the initial period after labor market entry affects their job mobility in later stages of the career. Third, to evaluate the empirical content of our hypotheses, we estimate the structural parameters of the model using data from the National Longitudinal Survey of Youth 1979 ($NLSY$ 79) covering the period between 1979 and 2006. Using the estimated parameters, we perform an empirical validation in which we compare the model's predictions about selected individual and aggregate labor market outcomes with their empirical counterparts. In what follows, we provide a brief sketch of each of the three parts of the paper and present some of the main results.

Using data from the $NLSY$ 79 we establish a number of empirical facts on individual job mobility in the U.S. We focus on white male labor force participants, mainly for homogeneity reasons. The main facts are as follows:

- The probability of separating from a new job declines significantly with age: 60% of all new jobs among U.S. white males aged $18-20$ years end within the first year. For workers aged $30-34$ years, this number is 40% and individuals aged 38 years and over face a 23% likelihood of separating from a new job within a year.

- The survival probability of new jobs drops when we consider a longer retention period: workers aged $18-20$ years have a 12% (7%) probability of retaining a new job for at least three (five) years, for individuals aged 38 years and over this number is 45% (36%).

- Job mobility is highest when workers start their careers, in later stages job changes become substantially less frequent. At the age of 45 years a typical U.S. white male has hold, on average, around 9 full time jobs, 50% of which are held within the first 5 years after labor market entry and 75% within the first 10 years.

- There is substantial variation in the total number of jobs across individuals. Some individuals will hold very few jobs in their entire career whereas others are change jobs frequently. All workers, however, share the common characteristic that the turnover uniformly declines along the career path.

- Unemployment is extremely high for individuals aged $18-20$ years but it falls sharply, within the age interval $18-25$ years, from an initial rate of 25%, down to roughly 8%. It continues to decline, up to age 35 where it levels out at a rate of 3%. For the remaining part of the working life cycle, that considered here, unemployment stays low and fluctuates in a narrow range between $2\% - 3\%$[1].

The evidence presented here hints towards an enormous amount of job mobility and turnover for individuals early in their career. In later stages, job attachments become substantially more durable and job changes less frequent.

To shed light on this issue we construct and estimate a life-cycle model of the labor market whose main characteristic is an information imperfection in the matching process. We model the imperfection as a noisy signal about the match quality that a firm and worker observe when they first meet. The key ingredient is that the imperfection is

[1]In the work presented here, we observe each individual up to a maximum age of around 47 years. Thus, we can not capture the unemployment dynamics before retirement.

assumed to be worker-specific and in particular it is linked to an individual's previous labor market history. Thus, the informativeness of the signal differs across workers as each of them has built up an idiosyncratic labor market history. The intuition is that an individual's labor market history seems likely to convey information about the worker's productive ability in a given job. The more information is available about past jobs the easier it is for the employer to screen the worker and to assess the suitability for a particular job. Therefore, young workers are typically associated with less informative signals than experienced individuals which have already built up an extensive history[2].

In the presence of imperfect information, the true quality of a match can turn out to differ from the employers ex-ante perception. Hence, employers occasionally make mistakes by matching with the wrong workers. The true match quality is an experience good about which agents learn, over time, based on observed output. Thereby, mismatches are detected and eventually dissolve. The survival probability of a given match is endogenously determined as it is conditional on the accuracy of the original signal. Given an informative signal, the true quality rather corresponds to the agents' ex-ante perception, whereas, with noisy information employers are more likely to be wrong. As a result, we observe different patterns of match separation and job mobility across agents that differ from each other in their labor market history and, in particular, in the amount of experience they have accumulated. The life cycle dimension is an important ingredient in this model. It allows us to capture the transition process through which young workers mature - by accumulating labor market experience - and, thereby, gradually move out of the initial state that is characterized by high job mobility.

We estimate the structural parameters of the model by indirect inference using data from the *NLSY* 79. Each individual in the *NLSY* 79 is observed since the point in time when it first enters the labor market until to date (conditional on not dropping out). This allows us to construct the entire labor market history of each individual using the information on the individual's actual labor market transitions. The empirical validation of the model delivers the following results:

- The estimated model can account very well, qualitatively and also quantitatively, for the observed pattern of job mobility along the life cycle. In line with the data, the model predicts that (a) young individuals are substantially more likely to separate from a new job than experienced workers, (b) the hazard rate of separating from a

[2] An alternative interpretation is as follows. One might think that the lack of experience causes young individuals to have an inaccurate perception about a particular job, which is expressed by a less informative signal. On the other hand, experienced workers can better assess, ex-ante, the suitability of a given job based on their experience accumulated on past jobs.

new job is increasing in the length of the retention period and (c) the hazard rate of separating from any given job is declining with tenure, for all age groups. Moreover, the model matches the observed extent of job turnover early in the career as well as the gradual decline in later stages. The implied path for the cumulated number of jobs, at each age, as a fraction of the career total is consistent with what we observe in the data.

- In the quantitative part, we distinguish between two worker types which can be considered as representing young and experienced workers respectively. The model allows some of the structural parameters, i.e. those that are governing the signaling process and, thus, determining the information imperfection, to differ across those two worker types. By how much they actually differ from each other is uncovered by the estimation. We find that the information imperfection, that is integrated into the matching process, is clearly worker-type-specific. Signals are fairly precise for individuals with a great deal of experience, whereas the average labor market entrant is associated with relatively noisy information. This is reflected by a signal to noise ratio which, the model predicts to be, roughly 4.5 times higher for experienced workers.

- The fundamental question, however, is for how long a labor market entrant has to stay in the market until the upgrade from the high to the low noise type is realized? Or put differently, how much labor market experience does a worker have to accumulate in order to be considered as a low noise type? The model estimates that the cumulated number of years of tenure required for a type change to occur is equal to 5.25.

- As mentioned previously, the quality of a given match is an experience good, about which agents learn over time. Learning, is assumed to be "all-or-nothing" . I.e. each period there is a constant probability that the match quality is either fully revealed or nothing is learned. We estimate this probability and find that it takes, on average, 1.62 years until the firm (and the worker) learns the true quality of a given match.

- The question that remains is how well the worker-type-specificity of informational frictions, that we incorporate, can account for the observed life-cylce pattern of job mobility. To address this issue, we consider a variant of the original model which allows also other factors, that potentially affect job mobility, to differ across worker types, namely (a) the job quality and (b) the exogenous rate of job destruction. We, then, use the estimates of the augmented model to disentangle the effect associated with each of the factors, and to measure their respective contribution in explaining the observed pattern. The statistic we use in this exercise, to compare the model to the data, is the one-year job retention rate for newly employed workers[3]. We find that the full model can account for 98.5% of the total observed increase in the retention rate over the life cycle. The informational frictions alone, can explain a sizable fraction of this increase, namely 41.8%. Allowing also the (true but initially unobserved) job quality to systematically differ across worker types adds another 45.5% to the explanatory power of the model[4]. The residual 11.2% can be explained by letting also the rate of exogenous job destruction to be type-specific.

The question at the core of this paper is: Why are young individuals more likely to separate from a new job than experienced workers? In the work presented here, we focus primarily on the role of informational frictions, and we find that they can explain a sizable fraction. Apart from the information-channel, there are certainly also other important factors that shape the observed patterns. Here are some thoughts about those: One might think that the young have a higher outside option than prime-age worker. Young adults can often rely on the support of their parents in case of an adverse economic event, such as job loss and unemployment. Typically they, too, have less financial obligations regarding child care, the financing of housing, etc. All this is likely to affect their quit behavior as a higher outside option, generally, makes a job loss less "painful".

Another channel concerns agents' search behavior over the life cycle. We may presume that the job search of an individual at beginning of the career is rather "undirected". Given the lack of labor market experience, young individuals will have a rather inaccurate perception about which job is the right one. Hence, they consider many jobs as suitable and very likely the one they pick will be a mismatch. Over time, the search behavior becomes more directed as agents learn, on the basis of their performance on previous jobs, which jobs are better suited than others. As a result, the range of jobs, in which agents search, narrows down as the career progresses. Those jobs that are selected, in the end, are likely to be good matches. The relevance of this channel is investigated in a follow-up paper.

[3] The one-year job retention rate measures the probability that a newly employed worker retains her current job for at least one year. We focus on this measure since we believe that it captures best the observed age-specificity of job mobility.

[4] With type-specific job quality, we will observe one type of workers, say the young, to be in jobs that are, on average, worse than the jobs of experienced workers. Reasons that for those systematic differences could be factors related to agents' experience, such as (a) agents' search behavior: young search less directed than experienced and are, therefore, more likely to end up in a mismatch, or (b) the level of job-specific human capital: the lack of experience might lower productivity the jobs of young less productive.

A sound understanding of individuals' job mobility is key for the design and the implementation of labor market policies. Our results indicate a, potentially important, path dependency in the working life cycle of individuals. We find that the job situation a worker faces in later stages of the career is closely linked to the outcomes and the experience accumulated during the early years. Therefore, any labor market policy, that affects the working conditions of the young is likely to generate spillovers into the entire working life cycle. The final part of this paper contains further considerations in that direction.

The remainder of the paper is organized as follows. In Section 2 we present the empirical facts that motivate our analysis. Section 3 outlines the structural model that we estimate later on. In Section 4 we outline the computational strategy that is adopted to solve the full model. Sections 5 and 6, respectively, present the estimation strategy and the data that is used to estimate the structural model. Section 7 documents and discusses our estimation results. Section 8 closes the main part of the paper with a brief conclusion. An Appendix is added which contains some of the results, tables and figures that we refer to in the text.

2 Empirical Facts

This section uncovers facts that document the variation in individual labor market mobility across age groups. The outcomes we focus on are various measures of job stability and turnover and unemployment rates.

Individual job stability/turnover

In this section we focus on measure of job stability and job turnover. The goal is to establish facts which are meant to illustrate how these measures systematically differ across age cohorts. Farber et al. (1993) and Topel and Ward (1992) find that younger workers have substantially higher rates of job loss than do older workers. Job attachments after entry into the labor market are extremely fragile. Topel and Ward (1992) find that, in the U.S., two thirds of all new jobs among young workers end in the first year. However, the probability of job loss declines with experience and jobs become more durable as the career progresses[5]. Their results imply that a worker with 12 years of experience is about half as likely to leave a new job as is a new entrant. Hall (1982) arrives at a similar conclusion. He estimates that the 5-years job retention probability for male workers aged $15-24$ years is only 3.93% whereas the rate for more experienced workers is substantially

[5]Notice that the probability of job loss in this context refers to the probability of loosing a *new* job.

higher. For workers aged $35-39$ ($45-49$) years this probability is equal to 16.0% (20.0%).

The job retention probability is a useful concept to measure job stability. It measures the probability that a worker with a given age and tenure will retain his current job for a certain number of years. Figure (1) and Table (1) report job retention probabilities calculated from $NLSY$ 79 data.

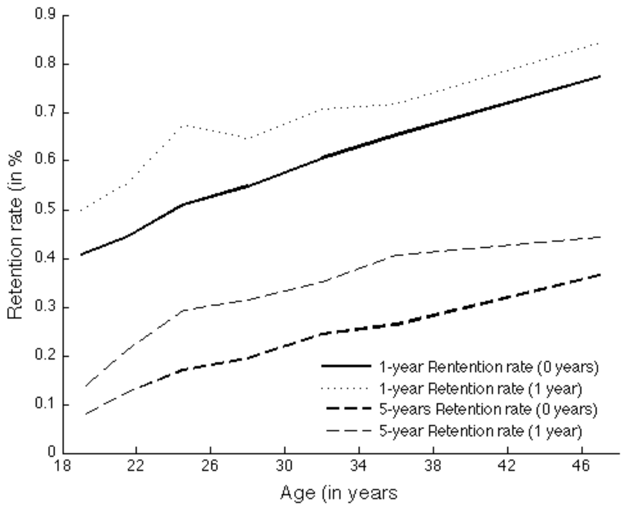


Figure 1

Comparison of the 1-year and 5-years job retention rates for workers aged 18-46 years with 0 years and 1 year of tenure. Source: Own calculations using data from the $NLSY$ 79

The sample considered here, consists of U.S. white male workers. Individuals attending school or serving in the army are excluded. Moreover, we consider only full time jobs - i.e. jobs with ≥ 30 weekly hours. The panel dimension of the $NLSY$ 79 allows us to construct the entire labor market history of each individual in the sample. Therefore, the reported retention rates are based on the actual labor market transitions of each observed individual. In particular, the rates are calculated from the number of workers in one age-tenure category who move on to higher age-tenure categories. For example, to compute the probability that a worker aged a years who has been on the job for k years will remain on the job for s more years, we use

$$\frac{\text{Number of workers aged } a+s \text{ years with } k+s \text{ years of tenure}}{\text{Number of workers aged } a \text{ years with } k \text{ years of tenure}} \tag{1}$$

Figure (1) depicts the $s \in \{1,5\}$-year job retention rates for workers aged $a \in \{18,46\}$ years with tenure $k = \{0,1\}$, that is I consider newly employed workers and those with one year of tenure. Three things are worth noting: (1) The job retention rate for workers with one year of tenure is always higher than the corresponding rate for newly employed workers. In other words, tenure substantially increases the likelihood of retaining the current job. This is not surprising as it is well know that the hazard rate of separating from an employer declines with tenure. This relation is also reflected here. (2) Also as expected, the 5-years retention rate is always lower than 1-year rate. The longer is the retention period the more likely it is that the job gets destroyed. (3) Most importantly, all reported job retention rates exhibit a strong increase with age. Evidently, young cohorts are much more likely to leave a new job than experienced workers. This pattern holds for all tenure classes and for all retention horizons.

Duration / Age	[18-20)	[23-26)	[30-34)	[38-
1 year	0.41 (0.49)	0.51 (0.67)	0.60 (0.71)	0.77 (0.84)
2 years	0.21 (0.31)	0.33 (0.51)	0.42 (0.57)	0.61 (0.58)
3 years	0.12 (0.21)	0.26 (0.42)	0.33 (0.43)	0.45 (0.56)
4 years	0.09 (0.15)	0.21 (0.33)	0.27 (0.39)	0.41 (0.47)
5 years	0.07 (0.13)	0.17 (0.29)	0.24 (0.35)	0.36 (0.44)

Job retention probabilities for workers with a given age (at the time of recruitment). Rows represent different job retention periods. Numbers in (without) brackets are the rates for workers with 52-103 (1-51) weeks of tenure. Source: Own calculations using data from the $NLSY$ 79

Table 1

The numbers underlying Figure (1) are reported in Table (1). The numbers in brackets are the corresponding retention rates for workers with one year of tenure. A worker aged $18-20$ years starting a new job has a 40% probability of retaining this job for at least 1 year. For a $30-34$ years old this probability is already 60% and a worker with roughly 20 years of labor market experience has a likelihood of 77% of retaining a new job for one year or longer. The retention probability drops significantly when we consider a longer retention horizon. For young age cohorts the probability of keeping a new job for at least three (five) years drops to 12% (7%) whereas the corresponding rates for experienced workers is 45% (36%). The numbers reported here are slightly higher than those found by Hall (1982). This is not surprising since the sample of workers considered here consists just of white males whereas Hall (1982) also includes females and blacks. Two groups which are known to feature less stable job attachments. To sum up, young workers when finding a job face a much higher probability of loosing this job within a relatively short period of time.

Evidently, mismatch is a phenomenon that is strongly concentrated among young cohorts. High rates of job loss for the young imply a lot of job turnover early in the career. A phenomenon which is also known as job-shopping. According to Topel and Ward (1992) a typical US male worker will hold seven full-time jobs which is about two thirds of his career total during the first 10 years in the labor market[6]. Own calculations, using the aforementioned *NLSY* 79 data, confirm these results. The first row in Table (2) depicts the mean number of (full-time) jobs for individuals with a given age. Standard deviations are in parentheses.

Age	[18-20)	[20-23)	[23-26)	[26-30)	[30-34)	[34-38)	[38-
				Total number of lifetime jobs			
	2.58	3.84	5.18	6.61	7.42	8.36	9.18
	(1.68)	(2.36)	(3.01)	(3.78)	(4.31)	(4.88)	(5.36)
				Number of jobs as a fraction of total number			
	0.332	0.484	0.629	0.746	0.806	0.882	0.963
	(0.211)	(0.236)	(0.232)	(0.211)	(0.201)	(0.165)	(0.084)

Upper panel: Mean number of (full-time) jobs for individuals until a given age. Lower panel: Mean number of jobs held until a given age, as a fraction of the career total. Standard deviations are in parentheses. Source: Own calculations using data from the *NLSY* 79

Table 2

Up to age 45 years a typical US white male worker has hold, on average, $8 - 9$ full time jobs. Notice that the "accumulation" of jobs over the life cycle is far from being uniform. The lower panel of Table (2) reports the number of jobs held, until a given age, as a fraction of the career total. 50% of the total jobs are held within the first 5 years after labor market entry and 75% of the total number is hold within the first 10 years (assuming that labor market entry occurs at age 18). This hints to an enormous amount of job turnover among young individuals. In later stages of the career job changes become substantially less frequent. Table (2) illustrates that the increase in the number of jobs flattens out at the age of roughly 30 years.

Notice that there is substantial variation in the total number of jobs across individuals. The standard deviation associated with the average number of jobs is strikingly high. This hints towards a large degree of heterogeneity regarding job changes. Some individuals will hold very few jobs in their entire career whereas others change jobs frequently. All workers, however, share the common characteristic that the turnover uniformly declines along the career path. This can be seen from the standard deviations reported in the lower Panel of Table (2). The turnover pattern we have identified, does apply not just to the U.S. labor market but it is a common characteristic of advanced OECD labor markets. Evidence for other countries, exists for the UK and for Germany. Booth, Francesconi and Garcia-Serrano (1997) report that half of the average five UK job changes occur in the

[6] The sample in Topel and Ward (1992) is taken from the Longitudinal-Employer-Employee-Data (LEED) which is based on individuals' Social Security earnings records.

first ten years. For Germany, Winkelmann (1997) finds that almost half of the average four German job changes fall in this period.

High rates of job turnover for young cohorts imply that the average tenure on each of the jobs is low. Table (3) depicts the average tenure of currently employed workers and the average duration of matches that ended.

Age	[18-20)	[20-23)	[23-26)	[26-30)	[30-34)	[34-38)	[38-
				Average tenure			
	0.87	1.46	2.13	3.11	4.43	5.85	7.91
	(0.91)	(1.38)	(2.06)	(2.91)	(3.95)	(5.09)	(6.77)
				Average duration of dissolved matches			
	0.494	0.751	1.094	1.429	1.931	2.681	3.665
	(0.58)	(0.92)	(1.33)	(1.79)	(2.47)	(3.35)	(4.52)

Upper panel: Average tenure of currently employed workers with a given age. Lower panel: Average duration of matches that ended for workers with a given age. Standard deviations are in parentheses. Source: Own calculations using data from the $NLSY$ 79

Table 3

Naturally, the average tenure increases in age. This is what one would expects. At the same time, the average duration of dissolved matches is basically half of the duration of existing matches. This is interesting especially for prime- age cohorts as it hints towards a coexistence of stable life-time jobs and high turnover. To shed more light on this issue we calculate the fraction of workers within a given tenure class. The results are reported in Table (4).

Tenure / Age	[18-20)	[23-26)	[30-34)	[38-
$0-2$	0.8952	0.6062	0.3715	0.2394
$2-5$	0.1011	0.2733	0.2645	0.2134
$5-10$	0.0036	0.1189	0.2446	0.2156
>10	0	0.0017	0.1196	0.3316

Fraction of workers of a given age within a certain tenure class. Source: Own calculations using data from the $NLSY$ 79.

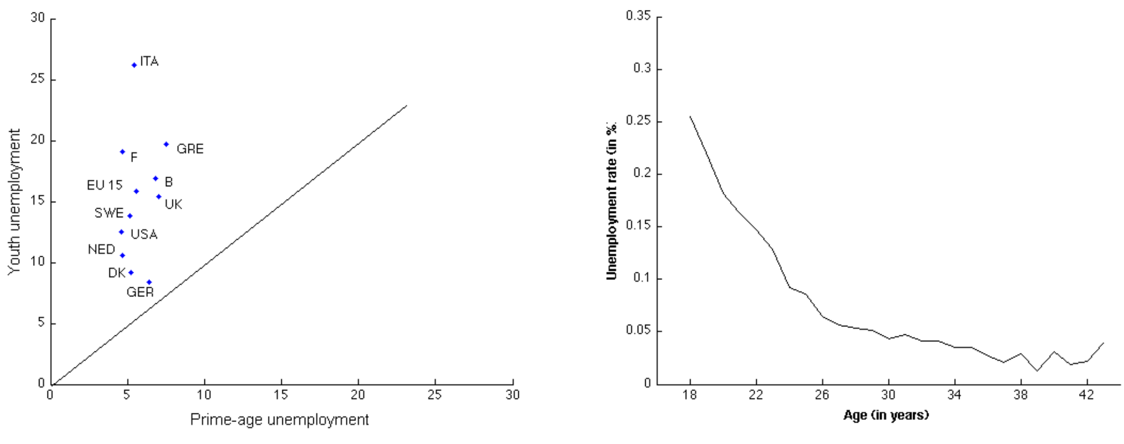
Table 4

For young cohorts the majority of the workers are in short-term jobs. This is implied by (a) the high job turnover of young workers and (b) the limited amount of time that has passed since labor market entry. When we consider later stages of the working life cycle, more and more individuals can be found holding medium and long-term jobs, though the fraction of short-term jobs among prime-age workers is still very high. For the cohort with age 38 years and higher one finds that there exists a strong dualism in the labor market. A substantial fraction of workers are in very stable jobs but at the same time a high proportion of individuals is subject to high turnover holding many jobs only for a short period of time. 33% of workers aged 38 years or older are in jobs lasting for 10 years and more. At the same time, almost one quarter of workers in the same age cohort have less than 2 years of tenure.

The general pattern, regarding job mobility over the life cycle, that emerges is the following. Young workers are subject to high turnover and generally face a high probability of job loss as opposed to experienced workers that are less affected by mismatch and usually end up in more stable jobs. After labor market entry young individuals typically hold a number of very brief jobs in the first few years. Eventually one job turns out to be a good match and lasts several years.

Unemployment

Unemployment and job turnover are naturally linked to each other since a job separation is often followed by a spell of unemployment. The findings of the previous section, thus, suggest that we should observe a strong age dependency also for unemployment.



(a) (b)

Figure 2

Panel (a): Cross-country comparison of the relation between the unemployment rates for young workers, aged 18-24 years, and prime-age workers aged 24-54 years. Source: OECD Labor Force Statistics. Panel (b): Unemployment rate for U.S. white male worker aged 18-43 years, Source: Own calculations using data from the *NLSY* 79.

Panel (a) in Figure (2) shows the average unemployment rates for young and prime-age male workers for the period $1985-2004$ for selected OECD countries[7][8]. Each of the points represents the combination of young-age and prime-age unemployment rates for a particular country. If a point lies above the 45° degree line then the unemployment rate of young exceeds that of prime-age workers and vice versa. For all the countries that are considered here, we observe a large discrepancy in unemployment rates between both age groups. Low rates of unemployment for prime-age workers typically coexist with high rates for young cohorts. In addition there is substantial cross-country variation in the difference between both rates. In absolute terms, the difference ranges from a minimum of 1.97 percentage points in Germany to 20.71 percentage points in Italy. Most of the countries can be found in the range of $7-13$ percentage points.

Panel (b) in Figure (2) provides a more detailed view for the U.S. labor market. In particular it draws the unemployment rate for U.S. white males workers belonging to different age cohorts. The pattern that is revealed is the following. Unemployment is extremely high for individuals aged $18-20$ years but it falls sharply, within the age interval $18-25$ years, from an initial rate of 25%, down to roughly 8%. It continues to decline, up to age 35 where it levels out at a rate of 3%. For the remaining part of the working life cycle, that considered here, unemployment stays low and fluctuates in a narrow range between $2\%-3\%$[9].

3 The Model

In this section we construct a life-cycle model of the labor market that is used to gather insights on the pattern of individual job mobility which we have identified above. The building blocks of the model are (a) firms and deterministically aging workers, matching in the labor market, (b) an informational friction that distorts the hiring process and introduces uncertainty about the quality of a job/worker match, (c) employers, learning about the match quality over time and (d) workers, building up an idiosyncratic labor market history which affects the informational friction. In what follows, we discuss each of the elements one after the other.

[7] The term "prime-age" refers to workers aged $24-54$ years.

[8] We focus on male workers only for matters of comparability. Arguably, some of the variation across age groups is likely to be driven by differences in labor force participation between male and female workers. Hence, putting both groups together would distort the picture.

[9] Notice that due to the limited time horizon of the *NLSY* 79 we do not observe individuals until retirement.

Workers

Time is discrete and denoted by $t = 0, 1, 2, ...$ The economy is populated by a continuum of risk-neutral individuals that are facing a finite life cycle. Let $k \in \mathcal{K} = \{0, 1, ...K\}$ denote an individual's age. The aging process is deterministic, i.e. an individual aged k at time t will be of age $k+1$ in period $t+1$. Individuals face a certain probability of dying that is age dependent. Let by $\rho_k \in [0, 1]$ denote the conditional survival probability from age k to age $k + 1$. Individuals live up to a maximum of K periods. At age K agents die with probability one, i.e. $\rho_K = 0$.

Worker - employed or unemployed - are characterized by their "type" denoted by i, where $i \in \mathcal{I} = \{1, 2, ..., \overline{I}\}$. The type is meant to capture all the "ability-related information" that could be of value to an employer to, ex-ante, screen and categorize a worker and to assess her productive ability for a given job. We can think of this information as being based on things such as curricula vitae, recommendations, personal interviews, test scores and the like. If a worker is of a high i-type, then more specific information is available upon which an employer can condition her hiring decision. Instead, for a low i-type worker, there is little information available. Arguably, with little information about an applicant, it will be hard for the employer to draw inferences on the worker's productivity and the quality of a potential match. Or in other words, when facing low i-type workers it will be hard for an employer to distinguish stars from lemons. The informational friction that we consider here is based on exactly this logic, but we will be more specific on that shortly.

Clearly, the amount of information, that is available, is related to workers' labor market experience. The longer an individual is in the market, the more information will be available on the basis of which employers can judge a worker's qualification for a given job. Therefore, i does not remain fixed over an individual's working life cycle but it changes over time. We make the following assumptions about the law of motion of i:

- At the moment of labor market entry, the worker's type is equal to the lowest possible value, i.e. when $k = 0$ we have that $i = 1$.
- The law of motion of i is stochastic and is governed by a discrete valued Markov process with certain transition probabilities. Those probabilities depend on the respective labor market state of the worker. Let by $\mu^s_{i|i'}$ denote the probability that a worker with labor market state s experiences a change of his type from i to i' within the current period. The set of labor market states we consider are include $s = \{e\text{-employed}, u\text{-unemployed}, b\text{-match break-up}\}$.

Workers do not have access to a savings technology, hence they consume their entire income every period. Preferences over consumption are assumed to be representable by a standard time separable utility function of the form $E\left\{\sum_{t=1}^{\infty}\beta^{t-1}u(c_t)\right\}$ where β is the time discount factor and c_t represents an individual's level of consumption. Expectations are taken with respect to the stochastic processes that govern mortality and the matching and signaling processes in the labor market.

Firms

A firm is assumed to consist of just one job. A firm's job can either be vacant or filled. Notice that there is not ex-ante heterogeneity on the firm's side, i.e. before being matched with a worker, firms are all homogenous. In order to produce output a firm has to be matched with a worker. A firm/worker match can be of good or of bad quality. Good matches produce y^g whereas bad matches produce y^b, with $y^g > y^b$. The match quality is an experience good, meaning that it is unknown to the firm and the worker when they first meet. The probability that a match between a type i worker and a firm is of good quality is given by π with $\pi \sim N(\alpha, \sigma_\pi^2)$ where $\alpha, \sigma_\pi^2 \geq 0$. π itself is unknown. Instead, agents (the firm and the worker) observe is a noisy signal about π which is denoted by γ. More precisely the signal γ is assumed to consist of the true value π and a noise component ϵ_i, i.e. $\gamma \equiv \pi + \epsilon_i$. The noise component is assumed to be worker specific. In particular we make the following assumptions about ϵ

- $\epsilon_i \sim N(0, \sigma_{\epsilon,i}^2)$ where $\sigma_{\epsilon,i}^2 \geq 0\ \forall i$
- $\sigma_{\epsilon,i'}^2 \leq \sigma_{\epsilon,i}^2$ for $i' \geq i$

The latter condition implies that the signal of low i-type workers will contain, on average, more noise than the signal of high i-type workers. From the assumptions above it follows that $\gamma \sim N(\alpha, \sigma_\pi^2 + \sigma_{\epsilon,i}^2)$. The firm and the worker form beliefs about π on the basis of the observed signal γ. Let $\hat{\pi}$ denote the expected probability that the match is good conditional on having observed the signal γ. Given the normality assumptions about π and ϵ we can write $\hat{\pi}$ as

$$\hat{\pi} \equiv E(\pi|\gamma, i) = (1 - \eta_i)\alpha + \eta_i\gamma \tag{2}$$

where $\eta_i = \frac{\sigma_\pi^2}{\sigma_\pi^2 + \sigma_{\epsilon,i}^2}$. It follows that $\hat{\pi} \sim N(\alpha, s_i^2)$ where $s_i^2 = \sigma_\pi^2\eta_i$. Notice that the worker's type is key for the formation of beliefs. The higher is the type of a worker the smaller is the variance of the noise component $\sigma_{\epsilon,i}^2$ and, hence, the signal will contain more information about the true match quality. Therefore more weight will be put on the signal γ. On the other hand, if $\sigma_{\epsilon,i}^2$ is high - meaning that the signal will be rather uninformative - more weight will be put on the mean α. In the extreme case of a totally uninformative signal - $\lim_{\sigma^2_{\epsilon,i}\to\infty} \hat{\pi} = \alpha$.

Match Formation

When a type i worker meets a firm with an unfilled job, both parties observe the signal γ and form beliefs about π. Those beliefs are expressed by the expected conditional probability $\hat{\pi}$. Conditional on the value of $\hat{\pi}$ both parties decide whether or not to form a match. When the match is not formed both continue their search. Instead, when the match is formed, the firm starts producing output in the subsequent period. The state of a firm with a filled job is, therefore, given by the triplet $(k, i, \hat{\pi})$ consisting of the age (k) and the type (i) of the employed worker and the common beliefs $\hat{\pi}$.

The output of a match depends on the match quality. In matches of unknown quality the firm and the worker don't observe the true output directly but rather $y = \overline{y} + e$, where $\overline{y} \in \{y^b, y^g\}$. If output was perfectly observable all the uncertainty about the match quality would vanish immediately after the match formation. We assume that the realizations of the output noise e are drawn from a uniform distribution with support $[-\mu, \mu]$. At first, the quality of a match is unknown. However, workers and firms learn about it over time on the basis of observed output. Given the assumptions on e, the learning process will be "all-or-nothing". This approach of modeling learning the match quality was first proposed by Pries and Rogerson (2005).

The name "all-or-nothing" refers to the process of how new information, that is generated within a period, is processed by the agents. In particular, if the realization of current output is such that $y > y^b + \mu$ $(y < y^g - \mu)$ the agents can infer that $\bar{y} = y^g$ $(\bar{y} = y^b)$. This implies that **all** the uncertainty about the match quality is removed at once. On the other hand if $y^g - \mu \leq y \leq y^b + \mu$ the agents have learnt **nothing** about the truth. The probability that the match quality is revealed within a period is, therefore, given by $\varphi = \frac{y^g - y^b}{2\mu}$ which is constant over time and independent of tenure. When a match is found to be bad it dissolves. On the other hand, good matches remain intact until either the worker dies or the match is hit by an exogenous separation shock.

Value Functions

Unemployed Workers

Workers can be either employed or unemployed. There is no labor force participation or retirement choice in the model. Unemployed workers encounter firms with open vacancies at an exogenous Poisson rate given by $\tilde{\lambda}$. Let $\Gamma^u(k,i)$ denote the value of unemployment to an individual with age k and type i. We can write $\Gamma^u(k,i)$ as follows

$$\Gamma^u(k,i) = b + \beta\rho_k\tilde{\lambda}\left(\int_{\hat{\pi}\in[0,1]} \Theta_u(k',i,\hat{\pi})dG^i(\hat{\pi})\right) + \beta\rho_k(1-\tilde{\lambda})\sum_{i'\in I}\mu^u_{i|i'}\Gamma^u(k',i') \tag{3}$$

where $\beta \in [0,1]$ is the personal discount factor and ρ_k denotes the probability that the worker survives to the next period. The value of unemployment consists of three parts: The first part - given by $b \geq 0$ - is the flow value of unemployment. We can think of b as representing, for instance, the value of leisure, unemployment benefits and the like.

The second part is the expected discounted value of encountering a firm with an open vacancy. Upon meeting a firm, a signal γ is drawn on the basis of which agents form beliefs $\hat{\pi}$. There is a one-to-one mapping between γ and $\hat{\pi}$. Therefore, we can omit the signal in the value function and work directly with the corresponding beliefs $\hat{\pi}$. Let $G^i(\hat{\pi})$ denote the cumulative probability distribution function associated with $\hat{\pi}$. Notice that the distribution of beliefs is worker-specific as the variance of the noise component $\sigma^2_{\epsilon,i}$ is a function of the worker's type.

The realized value of $\hat{\pi}$ will be crucial for whether or not a match is created. In equilibrium there will be an endogenous threshold for $\hat{\pi}$ above (below) which agents find it optimal (not optimal) to form the match. We will be more explicit about the determination of this threshold shortly. If the draw is sufficiently good then the match is formed, i.e. the worker agrees to stay with the firm - and vice versa. The benefit of accepting a job offer - i.e. the value of employment - is given by $\Gamma^e(k',i,\hat{\pi})$. If the worker finds it optimal to reject the job offer she continues searching. In this case the expected value of unemployment for next period is $\sum_{i'\in I}\mu^u_{i|i'}\Gamma^u(k',i')$. Notice that next period's value takes account for the possibility of a change in the worker's type i. Thus, we can write the maximum value of meeting a firm as

$$\Theta_u(k',i,\hat{\pi}) = \max\left\{\Gamma^e(k',i,\hat{\pi}), \sum_{i'\in I}\mu^u_{i|i'}\Gamma^u(k',i')\right\} \tag{4}$$

The expected value of a meeting is obtained by integrating (4) over all possible realizations of the beliefs $\hat{\pi}$. The third part of expression (3) is the expected discounted value of remaining unemployed. This value is realized with probability $(1-\tilde{\lambda})$ - which is the likelihood that an unemployed worker does not encounter a firm with a vacant job.

Firms

The state of an existing match is given by the triplet $(k, i, \hat{\pi})$. Let $J(k, i, \hat{\pi})$ denote the value of a job to a firm that employs a type i / age k worker and has beliefs $\hat{\pi}$. We can write J as follows[10]

$$J(k, i, \hat{\pi}) = \max\left\{E(y|\hat{\pi}) - w(k, i, \hat{\pi}) + \beta(1-\sigma)\rho_k J^+(k, i, \hat{\pi}), 0\right\} \tag{5}$$

where

$$J^+(k, i, \hat{\pi}) = \sum_{i' \in \mathcal{I}} \mu^e_{i|i'} \left[(1-\varphi)J(k', i', \hat{\pi}') + \varphi\hat{\pi}J(k', i', 1)\right]$$

A firm's outside option is always the value of an unfilled vacancy, which in equilibrium is equal to zero. Therefore, the max operator represents the firm's optimal choice with respect to current period employment. The value of a job to a firm consists of the two parts: the instantaneous return and the continuation value. The latter is denoted by $J^+(k, i, \hat{\pi})$. The instantaneous return is given by the difference between expected output, $E(y|\hat{\pi}) = \hat{\pi}y^g + (1-\hat{\pi})y^b$, and the wage, w, that is payed to the worker. Wages are determined each period by bilateral Nash bargaining between the worker and the firm[11]. Notice that the output realization of the current period is observed after wages have been negotiated. We do not allow for the within-period re-negotiation of wages.

With probability $(1-\sigma)\rho_k$ the match survives to the next period. Notice that $\sigma \in [0, 1]$ is the probability that the match is hit by an exogenous separation shock. Three events can occur in-between two periods: (a) the match quality remains unknown - which happens with probability $(1-\varphi)$ and yields payoff $J(k', i', \hat{\pi})$, (b) the match quality it is fully revealed and found to be good. This happens with probability $\varphi\hat{\pi}$ and the associated value is $J(k', i', 1)$. Or (c) the match quality is fully revealed and found to be bad. This happens with probability $\varphi(1-\hat{\pi})$. Bad matches break up, hence the continuation value

[10]It is important to note that due to the all-or-nothing learning the agents' beliefs, $\hat{\pi}$, stay the same throughout the period in which the quality of a match is unknown. This would be different if we had used a Bayesian learning approach instead. There, any new information that becomes available leads to an update of agents' beliefs. Also note that when the match quality is revealed $\hat{\pi}$ takes on the value of either 1 or 0 depending on whether the match is good or bad.

[11]The process of wage bargaining is described in detail in Appendix B.

associated with case (c) is equal to zero. If the match survives the worker's type might change. This is taken account for by the transition probabilities given by $\mu^e_{i|i'}$.

Employed Workers

Similarly we can write the value of a job to a worker as

$$\Gamma^e(k,i,\hat{\pi}) = \max\left\{ w(k,i,\hat{\pi}) + \beta\rho_k\Gamma^+(k,i,\hat{\pi}), \sum_{i'\in\mathcal{I}} \mu^b_{i|i'}\Gamma^u(k,i') \right\} \tag{6}$$

Every period the worker is free to quit her current job and to leave the firm for unemployment. A quit and the subsequent transition into unemployment might cause a change in the worker's type. Hence, the worker's outside option is given by the expected value of being unemployed $\sum_{i'\in I}\mu^b_{i|i'}\Gamma^u(k,i')$. The value of employment consists of the one period return, i.e. the current period wage income w, and the continuation value, i.e. the expected value of staying with the current employer also next period, Γ^+. The latter part is defined as follows

$$\begin{aligned}\Gamma^+(k,i,\hat{\pi}) = (1-\sigma)\sum_{i'\in\mathcal{I}} \mu^e_{i|i'}\left((1-\varphi)\Gamma^e(k',i',\hat{\pi}) + \varphi\hat{\pi}\Gamma^e(k',i',1)\right) + \\ [\varphi(1-\sigma)(1-\hat{\pi})+\sigma]\sum_{i'\in\mathcal{I}}\mu^b_{i|i'}\Gamma^u(k',i')\end{aligned}$$

- With probability $(1-\varphi)$ nothing is learned about the match quality. The implied next period's value is $\Gamma^e(k',i',\hat{\pi})$.
- If the match quality is found to be good - which happens with probability $\varphi\hat{\pi}$ - the value of the job changes to $\Gamma^e(k',i',1)$. Given that the worker stays employed she might experience a type transition which is captured by $\mu^e_{i|i'}$. The same is true also for the previous case.
- If the match is revealed to be bad - which happens with probability $\varphi(1-\hat{\pi})$ - or hit by an exogenous separation shock - with probability σ - the worker transits to unemployment. The value of unemployment is given by $\Gamma^u(k',i')$. Due to the break-up the worker might experience a change in her type. This is captured by the transition probabilities $\mu^b_{i|i'}$.

Equilibrium

All decisions, including those about match formation and break-up, are taken jointly by the firm and the worker. Hence the criterion, used in the decision making process, is the joint surplus of a match. We can write it as

$$S(k,i,\hat{\pi}) = J(k,i,\hat{\pi}) + \Gamma^e(k,i,\hat{\pi}) - \sum_{i' \in \mathcal{I}} \mu^b_{i|i'} \Gamma^u(k,i')$$

The joint surplus S consists of the sum of the value of the job to the firm and the worker net of heir respective outside values. The outside option for the worker is given by the value of unemployment, for the firm it is the value of an open vacancy (which in equilibrium is equal to zero). One can use the value functions in (5) and (6) to obtain an explicit expression for S. We state this expression in Appendix B. As mentioned previously, the firm and the worker's decision whether or not to form a match depends on the value of their common beliefs $\hat{\pi}$. We can define the reservation belief as the value of $\hat{\pi}$ for which the firm and the worker are indifferent between creating and not creating the match. We denote this value by $\underline{\hat{\pi}}$. For a given (k,i), $\underline{\hat{\pi}}$ has to fulfill

$$S(k,i,\underline{\hat{\pi}}) \begin{cases} > 0, & \forall \quad \hat{\pi} > \underline{\hat{\pi}} \\ = 0, & \text{for} \quad \hat{\pi} = \underline{\hat{\pi}} \\ < 0, & \forall \quad \hat{\pi} < \underline{\hat{\pi}} \end{cases} \tag{7}$$

In what follows we state a number of propositions that characterize some important properties of $\underline{\hat{\pi}}(k,i)$. The first establishes the existence and uniqueness of $\underline{\hat{\pi}}$ for all age cohorts k and all worker types i.

Proposition 1 *For each pair (k,i) there exists a unique threshold belief $\underline{\hat{\pi}}(i,k) \in (0,1)$ if $y^g > b \geq y^b$ holds.*

The proof is provided in Appendix A. The intuition of which goes as follows. First we show that the surplus functions $S(\cdot)$ is linear in $\hat{\pi}$ which rules out the existence of multiple $\underline{\hat{\pi}}$ that are consistent with the conditions stated in (7). Then we establish conditions for which the joint surplus for the lowest possible value of beliefs, i.e. $\hat{\pi} = 0$ is always negative. A necessary and sufficient condition for this to hold is $b > y^b$. This is intuitive as for $\hat{\pi} = 0$ the match produces output equal to y^b every period and thus if $b > y^b$ then the value of employment is always lower than the value of unemployment which guarantees a negative match surplus. Lastly, we seek for conditions that ensure $S(\cdot,\cdot,1) > 0$. A necessary and sufficient condition is $y^g > b$. The intuition is as before. For $\hat{\pi} = 1$ the match produces output y^g with probability one. Hence, if $y^g > b$ then the value of employment always exceeds the value of unemployment which guarantees a positive match surplus. The conditions needed for existence and uniqueness of an equilibrium, at the same time also ensure that good matches persist and bad matches break-up.

Proposition 2 *For a given type i the threshold belief $\underline{\hat{\pi}}(i,k)$ is increasing in age k.*

For this proposition we do not provide a fully-fledged proof but rather the intuition. Notice first that, due to the Nash bargaining, we have $J(k,i,\hat{\pi}) = (1-e)S(k,i,\hat{\pi})$, where e denotes a firm's bargaining power. Hence all the properties underlying $J(\cdot)$ are directly applicable to $S(\cdot)$[12]. From inspecting (5) we can infer that $J(\cdot)$ is essentially given by the finite sum over all $K-k$ single period returns $E(y|\hat{\pi}) - w(k,i,\hat{\pi})$ weighted by the survival probabilities ρ_k. Each of single period elements is necessarily positive. Therefore, the longer is the horizon over which a firm can collect the profits, the larger will be the total sum $J(\cdot)$. If, however, a worker has just a few periods left in the labor market then it will be less profitable for a firm to match with that worker. $J(k,i,\hat{\pi})$ is decreasing in k, hence, if $\underline{\hat{\pi}}(i,k)$ is such that $J(k,i,\underline{\hat{\pi}}(i,k)) = 0$ it follows that $J(k+1,i,\underline{\hat{\pi}}(k,i)) < 0$. At the same time we know that $\frac{\partial J(k,i,\hat{\pi})}{\partial \hat{\pi}} > 0$ (see Proof of Proposition 1), therefore $\underline{\hat{\pi}}(k+1,i) > \underline{\hat{\pi}}(k,i)$ is necessary to establish $J(k+1,i,\underline{\hat{\pi}}(k+1,i)) = 0$. Economically speaking this means that firms need to be compensated with a higher surplus when matching with workers that are close to the exit age. This is achieved by raising the threshold belief $\underline{\hat{\pi}}$.

Proposition 3 *For a given age k the reservation belief $\underline{\hat{\pi}}$ is increasing in the type i*

We provide the intuition underlying this proposition a little further below (for a simpler case though). This says that workers that are associated with a more precise signal, face a relatively higher threshold value than workers with a noisy signal. Next we define how the equilibrium can be characterized.

Definition 1 *A stationary equilibrium of the model consists of the following objects*

1. *A collection of threshold beliefs $\{\underline{\hat{\pi}}(k,i)\}_{k\in\mathcal{K},i\in\mathcal{I}}$ that satisfy condition (7)*
2. *A constant mass of firm/worker matches $\Phi^+(k,i,\hat{\pi})$ for all possible combinations of $(k,i,\hat{\pi})$, and a constant mass of unemployed workers $\Phi^-(k,i)$ for all combinations of workers' type i and age k.*

[12]Notice that the total surplus of a match is shared between the worker and the firm according to a sharing rule which is the result of a Nash bargaining process. In the bargaining process the firm and the worker set the wage $w(i,\hat{\pi})$ in order to maximize the Nash product that is given by $[J(k,i,\hat{\pi})]^e \left[\Gamma^e(k,i,\hat{\pi}) - \sum_{i'\in\mathcal{I}} \mu^b_{i|i'} \Gamma^u(k,i')\right]^{1-e}$. This maximization problem yields the familiar first order condition $(1-e)J(k,i,\hat{\pi}) = e\left[\Gamma^e(k,i,\hat{\pi}) - \sum_{i'\in\mathcal{I}} \mu^b_{i|i'} \Gamma^u(k,i')\right]$. By combining the first order condition with the definition of the joint match surplus one gets the surplus sharing rule used in the text.

Before we proceed to study certain properties of the equilibrium and compute aggregate variables, we provide a brief overview of the within-period timing of the model

1. At the beginning of the period firms and workers that were matched also the period before observe the current state of the match $(k, i, \hat{\pi})$. They then decide whether to terminate or to continue the match. Workers whose matches were dissolved enter unemployment where they receive b.

2. Wages w are negotiated bilaterally between firms and workers.

3. Matches that continue, produce output y. Wages w get payed to workers.

4. On the basis of the current output realization the firm and the worker update their (common) beliefs about the true quality of the match. This is done in an all-or-nothing fashion. With probability φ the true match quality is revealed and with probability $1-\varphi$ nothing is learned. Matches that are found to be bad are dissolved.

5. Unemployed workers encounter firms with an open job with probability $\tilde{\lambda}$. Each pair that meets observes a signal γ about the quality of the match. Conditional on the signal they form beliefs about the true probability that the match is good. Those beliefs are expressed by $\hat{\pi} \equiv E(\pi|\gamma, i)$. On the basis of $\hat{\pi}$ they decide whether or not to form a match. New matches start producing output the period after.

6. At the end of the period workers' age and experience a change in their type i depending on their labor market state.

7. Workers die with probability $(1 - \rho_k)$, jobs get hit by the destruction shock with probability σ and new workers - aged $k = 0$ - enter the market.

4 Solving the Full Model

This section outlines the solution algorithm that was used to compute the equilibrium of the full model presented in Section 3. The state of an unemployed is given by (k, i) and that of a match is $(k, i, \hat{\pi})$. Notice that k (age) and i (type) are both discrete objects. The aging process in the model is constructed in a way so that an individual with age k today is of age $k' = k + 1$ next period. Furthermore, the process that governs the type transition is a discrete valued Markov process implying that the realizations of i all lie on a pre-specified grid. This suggests the use of equally-spaced grids for k and i which we specify as follows: $\mathcal{K} = \{1, 2, ..., \bar{K}\}$ and $\mathcal{I} = \{1, 2, ..., \bar{I}\}$. $\bar{K}$ is the maximum number of periods an individuals stays in the labor market and $\mathcal{I}$ is the highest achievable type. For the beliefs $\hat{\pi}$ we adopt a continuous representation of the state space.

The value function of an unemployed worker and the joint surplus function of a match are described by the mappings $\Gamma : \mathcal{K} \times \mathcal{I} \rightarrow \Re$ and $S : \mathcal{K} \times \mathcal{I} \times [0,1] \rightarrow \Re$. Due to the discreteness of its domain we can express Γ as a collection of points $\{\Gamma_{k,i}\}_{k\in\mathcal{K},i\in\mathcal{I}}$. S, on the other hand, needs to be approximated. We do not use one and the same method to approximate S on the entire $[0,1]$ domain but, instead, we combine two different methods. The reason for that is the kink in the value function that is induced by the cutoff value $\underline{\hat{\pi}}$. In particular we have $S(k,i,\hat{\pi}) = 0 \ \forall \hat{\pi} \leq \underline{\hat{\pi}}$ and $S(k,i,\hat{\pi}) > 0 \ \forall \hat{\pi} > \underline{\hat{\pi}}$. To accommodate for that we use a piecewise-linear approximation of S for values of $\hat{\pi} \in [0, \underline{\hat{\pi}})$ and a Chebyshev polynomial for $\hat{\pi} \in [\underline{\hat{\pi}}, 1]$. For each node on the $\mathcal{K} \times \mathcal{I}$ grid we construct the polynomial by running a Chebyshev regression using 50 collocation nodes. These nodes are given by the collection $\{z_{k,i,x}\}_{k\in\mathcal{K},i\in\mathcal{I},x\in\{1,2,\ldots,50\}}$.

The value functions Γ and S are found jointly using an algorithm that incorporates a fixed-point iterative scheme. The algorithm is structured in the following manner:

1. The first step involves guessing a joint surplus function S. The initial guess is denoted by S^0.

2. The second step involves computing the threshold beliefs. For each combination of (k,i) we identify the $\underline{\hat{\pi}}$ for which $S(k,i,\underline{\hat{\pi}}) = 0$.

3. For each pair (k,i) we allocate the collocation nodes in the interval $[\underline{\hat{\pi}}(k,i), 1]$ and run the Chebyshev regression to find the approximation coefficients associated with S. Those coefficients are given by $\{c_{k,i,x}\}_{k\in\mathcal{K},i\in\mathcal{I},x\in\{1,2,\ldots,50\}}$.

4. Next, we solve for $\{\Gamma_{k,i}\}_{k\in\mathcal{K},i\in\mathcal{I}}$ by first using the terminal condition $\Gamma(\bar{K}, i) = b$ for $\forall i$ and then working backwards. The integral in Equation (3) is computed using a Gauss-Legendre quadrature with 7 quadrature nodes.

5. To compute $\{S^g_{k,i}\}_{k\in\mathcal{K},i\in\mathcal{I}}$ it suffices to evaluate $S(k,i,\hat{\pi})$ at $\hat{\pi} = 1$, $\forall k,i$. This follows from $S^g(k,i) = S(k,i,1)$.

6. To solve for the new S^{j+1} we proceed as follows: We use (a) the value of the previous iteration step S^j, (b) the terminal condition $S(\bar{K}, i, z) = 0$, $\forall i, z$, (c) the value of unemployment $\{\Gamma_{k,i}\}_{k\in\mathcal{K},i\in\mathcal{I}}$ and work backwards for all (k,i,z).

7. Lastly, we evaluate $\|S^{j+1} - S^j\|$. If the distance is smaller than some convergence criterion we terminate the algorithm, else we set $S^j = S^{j+1}$ and return to step 2.

5 The Estimation Strategy

The structural model is estimated by indirect inference. This method, which was first proposed by Gourieroux, Montfort and Renault (1993), belongs to the class of simulation-based estimation procedures[13]. Estimators of this sort are particularly convenient in situations when the complexity of the model leads to an intractable likelihood function or when some of the variables are unobserved in the data. The idea of indirect inference as it is used here is to match the statistical properties of the simulated data with those of observed data along selected dimensions. The dimensions, along which the simulated and the observed data are evaluated, are represented by an auxiliary model. The central idea is then to choose the parameters of the structural model in a way so that the parameters of the auxiliary model estimated from actual and from simulated data are as close as possible.

Obviously, a crucial step in the estimation procedure concerns the selection of the auxiliary model. A "good" auxiliary model is one that captures well the statistical properties of the data along the dimension(s) required to identify the model's structural parameters. The parameters estimated from the auxiliary model have to contain enough information in the sense that they have to be sufficiently responsive to changes in the model's structural parameters. If that was not the case then the objective function would exhibit relatively little curvature and the structural parameters would be poorly identified. In this respect, the suitability of an auxiliary model can be judged by looking at the standard errors of the estimated of the structural parameters. Secondly, the auxiliary model should be fast to compute. The estimation of the structural parameters typically requires a high number of evaluations of the model's objective function. This can potentially be very time consuming. Hence, choosing a parsimonious auxiliary model is key for keeping the computational time at a reasonable level.

As the auxiliary model, I choose a discrete-time hazard model that is tailored to the analysis of duration data. The choice is motivated by the nature of the economic process that we are looking at. The structural model is designed to capture differences in job duration across age cohorts. Therefore, a hazard rate model is a natural candidate since it allows to identify the effect of individuals' characteristics on the actual job duration.

The data that is used in the estimation consists of observations of individuals' employment spells. Each observation is indexed by $j \in \{1, 2, ..., J\}$, where J denotes the

[13] Similar estimation methods include efficient method of moments (Gallant and Tauchen (1996)), or simulated (quasi-)maximum likelihood (Smith (1993)), For a discussion of simulation-based estimators see Tauchen (1997) and Gourieroux and Monfort (1996).

total number of observations. For the estimation I divide the time line into N intervals that are given by $[0, \tau_1)$, $[\tau_1, \tau_2)$, ..., $[\tau_{N-1}, \tau_N)$. Intervals are indexed by the subscript $n \in \{1, 2, ..., N\}$. Let t_j denote the duration of employment spell j and let $y_{j,n}$ be a binary variable that is zero if in interval n the worker is still employed in job j and unity otherwise. Thus, $y_{j,n}$ will be zero for all intervals n for which $\tau_n < t_j$ and it will be one for all intervals or which $\tau_n \geq t_i$. For each employment spell we observe a string of zeros followed by a string of ones[14]. The important information is the time interval for which $y_{j,n}$ becomes unity for the first time. This is because the switch from zero to one indicates a match separation.

Notice, however, that some observations are right-censored. That is, we can not determine the exact time when the job ended. This can be due to two reasons: (a) the duration of the employment spell t_j exceeds the time that is captured by the intervals - that is the case whenever the job lasted longer than the period that is captured by the N intervals, i.e. $t_j \geq \tau_N$ - or (b) if the employment spell is still ongoing at the time the latest survey was taken. This was in 2006. In the first case all observations are right-censored at τ_N, in the second case the observation is right-censored in the interval corresponding to the job duration at the time of the survey. To account for censoring we construct the binary variable c_j that is unity when the duration of job j is uncensored and zero otherwise. Notice that for the estimation it will be important whether the ending of an employment spell represents a true job loss or is due to censoring. This information will be captured by c_j.

The hazard rate of separating from job j in interval n is modeled as a piecewise constant function given by

$$\lambda(t; \mathbf{x}_j, \beta) = \kappa(\mathbf{x}_j, \eta)\lambda_n, \qquad t \in [\tau_{n-1}, \tau_n) \tag{8}$$

β is a vector of unknown parameters and $\kappa(\mathbf{x}_j, \eta) > 0$ is a function of observable worker characteristics. Those characteristics are captured by the vector $\mathbf{x}_j$, whereas η is a vector of unknown parameters. The baseline hazard for each interval is given by λ_n. This specification implies a constant hazard within a particular interval n, however it allows the hazard to differ across intervals. Notice that the existence of a baseline hazard makes a constant term as a covariate superfluous. For a particular employment spell j only the intervals for which $y_{j,n} = 0$ can be used to predict the probability of $y_{j,n+1}$ being zero or one in the subsequent interval $n + 1$[15]. These (conditional) probabilities are given by

[14] Notice that $y_{j,n} = 1$ implies $y_{j,n+1} = 1$.

[15] This is due to the fact that once $y_{j,n} = 1$ we know that $y_{j,n+1} = 1$.

$$
\begin{aligned}
P(y_{j,n}=1|y_{j,n-1}=0,\mathbf{x}_j) = &\ P(\tau_{n-1}\leq t_j<\tau_n|t_j\geq\tau_{n-1},\mathbf{x}_j)\\
= &\ 1-\exp\left[-\textstyle\int_{\tau_{n-1}}^{\tau_n}\lambda(s;\mathbf{x},\beta)ds\right]\\
= &\ 1-\exp\left[-\kappa(\mathbf{x}_j,\eta)\lambda_n\left(\tau_n-\tau_{n-1}\right)\right]
\end{aligned} \tag{9}
$$

$$
P(y_{j,n}=0|y_{j,n-1}=0,\mathbf{x}_j) = \exp\left[-\kappa(\mathbf{x}_j,\eta)\lambda_n\left(\tau_n-\tau_{n-1}\right)\right]
$$

Using those probabilities one can construct the log likelihood function. If for employment spell j the separation - censored or uncensored - occured in interval n_j the log likelihood is given by

$$
\begin{aligned}
l(n_j,\mathbf{x}_j,c_j;\beta) = &\ -\textstyle\sum_{h=1}^{n_j-1}\kappa(\mathbf{x}_j,\eta)\lambda_h\left(\tau_h-\tau_{h-1}\right)\\
&+c_j\log\left\{1-\exp\left[-\kappa(\mathbf{x}_j,\eta)\lambda_{n_j}(\tau_{n_j}-\tau_{n_{j-1}})\right]\right\}
\end{aligned} \tag{10}
$$

The first part is the probability the job lasts until n_j and the second part is the probability that it ends in the interval n_j. Notice that the latter part is non-zero only when the observation is uncensored and therefore represents a true job loss. The log likelihood function for the entire sample is obtained by summing over $j=1,...,J$ employment spells. It is given by

$$
\mathcal{L}(\mathbf{y}_J;\beta)=\sum_{j=1}^{J}l(n_j,\mathbf{x}_j,c_j;\beta) \tag{11}
$$

where $\mathbf{y}_J=\{n_j,\mathbf{x}_j,c_j\}_{j=1}^{J}$

At this point let us discuss the worker characteristics that are considered and the functional form of $\kappa(\cdot,\cdot)$. For a particular employment spell j we use information on: (a) a_j: the worker's age at the time when the match was formed and (b) ω_j: the percentage deviation of the worker's initial wage in job j from the average initial wage payed in matches that were formed by workers of the same age. According to the structural model a_j and ω_j both contain important information about an individual's hazard rate. Recall that a worker's age at match formation is a good predictor of her labor market experience which in turn is indicative for the worker's type. The type determines the precision of the signal that is emitted in the hiring process and it thereby affects the survival probability of the employment relationship. A worker's age at match formation will, therefore, be informative about the likelihood of separation.

The channel through which ω_j affects the hazard rate goes via agents' beliefs. Agents are rational in the sense they have correct beliefs about the probability that the match is good[16]. Consequently, agents' beliefs are inversely related to the **actual** likelihood of match separation. The probability of survival determines the joint surplus of a match. More durable matches generate a higher surplus to the worker and the firm than matches which are expected to break-up soon. Typically the value of the job is reflected by the worker's remuneration, i.e. the wage. As a result, a positive deviation of the worker's initial wage from the average initial wage - reflected by a positive ω_j - is an indication that agents expect their employment relation to be stable and to yield a high return.

We do not treat the workers age at match formation as a continuous variable. Instead, we consider age intervals and work with dummy variables. In particular we divide the "age-line" into $K+1$ distinct intervals that are given by $[a_0, a_1)$, $[a_1, a_2)$, ..., $[a_{K-1}, a_K)$, $[a_K, \infty)$. Let $\tilde{a}_k$ be a dummy variable that is unity when the worker's age at match formation falls in the interval $[a_{k-1}, a_k)$, and zero otherwise. The vector of covariates $\mathbf{x}_j$ is thus given by $(\tilde{a}_1, ...\tilde{a}_K, \omega_j)$. As a functional form for $\kappa(\cdot, \cdot)$ we choose

$$\kappa(\mathbf{x}_j, \eta) = \exp\left(\eta_{a_0|a_1}\tilde{a}_1 + \eta_{a_1|a_2}\tilde{a}_2 + ... + \eta_{a_{K-1}|a_K}\tilde{a}_K + \eta_\omega \omega_j\right)$$

The age category that serves as a means of comparison is given by the interval $K+1$, i.e. $[a_K, \infty)$. Hence, any given age coefficient $\eta_{a_{k-1}|a_k}$ $k \in \{1, 2, ..., K\}$ has to be interpreted as the difference in the hazard between a worker that starts a new job at age $a \in [a_{k-1}, a_k)$ and a worker that starts a job at age $a \geq a_K$ - everything else equal. If those coefficients were all the same across age cohorts then, in terms of separation probability, it would not matter at what age a worker starts a new job.

Using the information on $(n_j, \mathbf{x}_j, c_j)$ that is available for each employment spell j we can estimate the coefficients of the auxiliary model by maximizing the log likelihood function in (11). The vector of coefficients we estimate are given by the vector $\beta = (\eta_{a_0|a_1}, ..., \eta_{a_{K-1}|a_K}, \eta_\omega, \lambda_1, ...\lambda_N)$. The maximum likelihood estimate of β is

$$\hat{\beta}_J = \arg\max_\beta \mathcal{L}(\mathbf{y}_J; \beta) \tag{12}$$

Recall that he ultimate goal of the estimation process is to determine the parameters of the structural model. Estimating the coefficients of the auxiliary model from data can be considered as an intermediate step which is a convenient way to condense the information that is contained in the data into something very tractable, i.e. the coefficients of the auxiliary model.

[16]The term "correct" should reflect that agents' beliefs are, on average, consistent with the true match quality.

We now turn to the next step in the estimation procedure. Let $\mathbf{y}_S(\phi)$ denote a matrix of S simulated observations of the endogenous variables $\{n_s, \mathbf{x}_s, c_s\}_{s=1}^S$ using the structural model and a set of structural parameters given by ϕ. Let $\beta(\mathbf{y}_S, \phi)$ denote the coefficients of the auxiliary model estimated from simulated data $\mathbf{y}_S(\phi)$. Estimating the parameters of the structural model essentially amounts to finding the value of ϕ so that the distance between the coefficients of the auxiliary model estimated from simulated data and from actual data is minimized. Technically, ϕ is chosen so that

$$\tilde{\phi}_S^H(\mathcal{Q}) = \arg\min_{\phi} \left(\hat{\beta}_J - \frac{1}{H}\sum_{h=1}^{H}\beta^h(\mathbf{y}_S, \phi)\right) \times \mathcal{Q} \times \left(\hat{\beta}_J - \frac{1}{H}\sum_{h=1}^{H}\beta^h(\mathbf{y}_S, \phi)\right)' \tag{13}$$

where $\mathcal{Q}$ is a symmetric, non-negative definite weighting matrix. Gourieroux, Montfort and Renault (1993, Proposition 4) show that the optimal weighting matrix in this case is given by $\mathcal{Q}^* = \mathcal{Z}_0^{-1}$ where

$$\mathcal{Z}_0 = \lim_{J\to\infty} V\left[\sqrt{J}\frac{\partial\mathcal{L}(\mathbf{y}_J, \beta_0)}{\partial\beta}\right]$$

Gourieroux, et al. (1993) show that under usual regularity conditions the estimator $\tilde{\phi}_S^H(\mathcal{Q})$ is asymptotically normal, when H is fixed and S goes to infinity:

$$\sqrt{S}\left(\tilde{\phi}_S^H(\mathcal{Q}^*) - \phi_0\right) \to^d N\left[0, \mathcal{W}(H, \mathcal{Q}^*)\right]$$

where the variance-covariance matrix $\mathcal{W}(H, \mathcal{Q}^*)$ is given by

$$\mathcal{W}(H, \mathcal{Q}^*) = \left(1 + \frac{1}{H}\right)\left[\mathcal{Z}_1'\mathcal{Q}^*\mathcal{Z}_1\right]^{-1} \tag{14}$$

The matrices $\mathcal{Z}_0$ and $\mathcal{Z}_1$ are impossible to compute. However, Gourieroux, et al. (1993) show that both can be consistently estimated by evaluating the Hessian using, respectively, observed and simulated data.

$$\mathcal{Z}_0 = p\lim_J J\frac{\partial^2\mathcal{L}(\mathbf{y}_J; \hat{\beta}_J)}{\partial\beta'\partial\beta} \qquad \mathcal{Z}_1 = p\lim_S \frac{\partial^2\mathcal{L}(\mathbf{y}_S(\tilde{\phi}_S^H(\mathcal{Q}^*)); \hat{\beta}_J)}{\partial\beta'\partial\phi} \tag{15}$$

6 The Data

The data I use in the estimation comes from the National Longitudinal Survey of Youth 1979 (*NLSY* 79). The *NLSY* 79 is a representative sample of 12,686 young men and women who were 14 to 22 years of age when first surveyed in 1979. The cohort was interviewed annually through 1994. Since 1994, the survey has been administered biennially. The last survey that is available was published in 2006. The entire data set I use in the estimation thus consists of 27 individual surveys. The *NLSY* 79 contains demographic variables, labor market data and information on individuals' wealth and consumption.

The advantageous feature of the *NLSY* 79 is it's panel dimension. Each individual in the survey is observed since the point in time when it first enters the labor market until to date (conditional on not dropping out). This allows me to construct the entire labor market history of an individual using the information on the individual's labor market transitions. It might seem problematic though that the survey is taken annually (and from 1994 even biennially) considering that a substantial fraction of jobs - especially for young individuals - lasts for less than a year. However, the survey is constructed in a way so that during each interview the individual is asked not just about any ongoing job but also about up to 5 employment spells that occurred since the last interview. Therefore, it is possible to recover the complete labor market history of each individual in the survey.

The data set I construct consists of employment spells of individuals, each denoted by j, and a set of individuals characteristics associated with the respective employment spell. As mentioned previously, each observation j consists of information about (a) t_j: the worker's total tenure on the job, (b) c_j: an indicator that is unity when the duration of job j is uncensored and zero otherwise, (c) a_j: the worker's age when the match was formed and (d) ω_j: the percentage deviation of the worker's initial wage in job j from the average initial wage payed in matches that were formed by workers of the same age. To guarantee a reasonable degree of homogeneity across individuals I exclude certain observations from the data set. In particular, the final data set consists of observations of white males that had their first job at the age of 18 or later. Moreover, I exclude part time jobs and consider just those employment spells for which individuals worked for $\geq$ 30 hours per week[17].

In constructing the data set and in particular to determine the length of an employment spell I take a retrospective approach. More precisely, every time I observe a job separation in the data I use the information on the worker's total tenure on that job to determine the length of the employment spell. Individuals are always asked to report the

[17]The decision to exclude part-time jobs is based on the following consideration. It seems likely that firms' hiring practices and workers' job selection decisions are based on very different criteria depending on whether a part-time or a full-time job is concerned. Full-time jobs naturally involve a higher degree of commitment from both sides and hence the decision to form a match might differ between both types of jobs.

total number of weeks they have been working for a particular employer - irrespective of whether the job has already ended or not. Therefore, at the time of the separation I read off the worker's tenure on the job to determine the duration of the job.[18]. For the jobs that are still ongoing at the time of the last survey (2006) we do not observe a separation. Therefore, I record the duration of the job at the time of the last survey and mark the observation as right censored.

The next step concerns the construction of a variable that - for each employment spell j - indicates the worker's age at the time of the match formation. To that end I use information on the worker's age at the time of the job separation and the worker's total tenure on the job. However, using simply the worker's age measured in years which is reported during the interview is not advisable. This is mainly because, for our purposes a "year" as a unit of measurement is too coarse to capture the worker's exact age at the time of the match formation. To minimize the error I, thus, compute the worker's exact age (in month) at the time of the interview. This is done by using information on the month and the year of the current interview and the month and the year of the individual's birth. By subtracting the worker's tenure on a particular job (measured in months) from her current age one gets the individual's precise age when she started the job.

The last step concerns the construction of a variable which measures the percentage deviation of the worker's initial wage in job j from the average initial wage payed in all matches that were formed by workers of the same age. The construction of this variable is a little tricky mainly because the initial wage needs to be recovered for each employment spell individually. To this end I start, for each employment spell, at the time of separation. In step (1), I use information on (a) the week number of the current interview, (b) the week number of the last interview and (c) the total tenure of worker on the current job. By using (a) and (b) we can determine how many weeks have passed since the last interview. In step (2), I consult the preceding survey and check whether the very same job has a duration of less or equal than 52 weeks. If that is the case I record the reported wage as the worker's initial wage in that job. If the duration is more than 52 I return to step (1) and continue until the initial wage is found.

[18] The two main advantages of this approach are that (a) I can determine the exact length of each employment spell and (b) I avoid the possible double-counting of employment spells. An alternative approach would be to go the other way around, i.e. to follow a job from the beginning to its end and to record the tenure at each consecutive survey for which the job still exists. This, however, would be very fault-prone, because any temporary interruption of an individual's record due to the non-availability for the interview, for instance, would mistakenly be considered as the ending of an employment spell. The double counting of a single spell would be unavoidable.

This method has two main advantages: first it can easily handle the switch from annual to biennale interviews without loosing information. This is because it uses the number of weeks between two consecutive interviews to identify the survey in which one has to look for the respective job. Second, for the very same reason it can manage a temporary discontinuation of the observability of a job which might be due to the non-availability of the interviewee.

Number of employment spells, by duration (in years) and by age at match formation								
Age / Duration	[0-0.5)	[0.5-1)	[1-2)	[2-3.5)	[3.5-5)	[5-10)	[10-	Total (age)
[18-20)	996	439	367	163	87	87	36	2,172
[20-23)	2,286	975	806	443	211	218	171	5,110
[23-26)	1,996	985	866	456	205	241	230	4,979
[26-30)	1,591	954	839	437	215	256	320	4,612
[30-34)	885	563	541	340	164	237	302	3,032
[34-38)	461	318	397	309	164	351	114	2,114
[38-	633	478	692	525	300	340	20	2,988
Total (duration)	8,848	4,712	4,508	2,673	1,346	1,727	1,193	25,007

Number of observations across tenure and age groups. "Age" refers to the workers age at the time of recruitment. Sample consists of white-male U.S. workers. Data: *NLSY* 79

Table 5: Summary Statistics I

Moreover, it is fairly precise. For 79.14% of the employment spells I am able to recover the initial wage. The failure to get a complete assignment of initial wages to jobs is due to a variety of reasons, the most important of which is the existence of missing values for wages. By inspecting the original data it becomes evident that individuals very often do not report their wages especially for jobs with a short duration. Arguably, excluding those observations with missing values from the data set would therefore introduce a bias. To solve this problem I choose to use imputed values to find the remaining initial wages. The remaining steps in constructing ω_j are straightforward. For each age cohort I compute the average initial wage and ω_j is then simply the percentage deviation of the initial wage of each observation from the mean.

The resulting data set consists of $25,007$ observations of employment spells. 90.61% of which are uncensored. Table (5) depicts in detail how the total number of observations is distributed among the different tenure and age groups. Notice again that the term "age" refers to the workers age at the time when the match was created, and not to the worker's current age. Not surprisingly, the main bulk of employment spells in the data set is of short duration. The proportion of jobs surviving into higher tenure classes shrinks as duration increases. Hence, the distribution of jobs across duration is strongly left skewed. This pattern holds throughout all age cohorts. Notice, however, that cohorts with higher age, generally, have a larger proportion of medium- and long-term jobs than young cohorts.

Job duration	[0-0.5)	[0.5-1)	[1-2)	[2-3.5)	[3.5-5)	[5-10)
Mean age	25.78	27.01	28.32	29.37	29.78	30.32
(Std.dev. Age)	(6.28)	(6.81)	(7.57)	(7.65)	(7.71)	(6.98)
Mean ω	-0.105	-0.032	0.043	0.079	0.127	0.136
(Std.dev. ω)	(0.35)	(0.42)	(0.48)	(0.51)	(0.57)	(0.56)
Initial age	**[18-20)**	**[20-23)**	**[23-26)**	**[26-30)**	**[30-34)**	**[34-38)**
Mean tenure	1.435	1.821	2.082	2.444	2.913	3.003
(Std.dev tenure)	(2.87)	(3.85)	(4.05)	(4.21)	(4.18)	(3.29)
Mean ω	0.00	0.00	0.00	0.00	0.00	0.00
(Std.dev. ω)	(0.30)	(0.33)	(0.39)	(0.48)	(0.51)	(0.59)

Upper panel: Mean and the standard deviation of workers' age across tenure classes (first two rows). Mean and the standard deviation of ω across tenure classes (third and fourth row). Lower panel: Mean job duration across age groups (fifth row). The standard deviation of ω (last row). Sample consists of white-male U.S. workers. Data: $NLSY$ 79

Table 6: Summary Statistics II

Table (6) provides summary statistics about the worker characteristics. The first two rows in the upper panel depict the mean and the standard deviation of workers' age across tenure classes. What we observe is the following pattern: There is a positive relationship between the duration of an employment spell and the average age at which an individual has started working in the job. Or in short: matches which last longer are, on average, created by higher age cohorts. Short-term jobs, i.e. those that end within half a year - are populated by workers aged on average 25.78 years, whereas long-term jobs - those that last up to 10 years - are created by workers aged on average 30.32 years. However, the positive relation that we observe is not particularly strong. The standard deviation of age is generally quite large across tenure classes.

The next two rows in the upper panel depict the mean and the standard deviation of ω across tenure classes. What we observe is in line with our theory. A central prediction of our model is that matches which pay above- average initial wages are likely to survive longer. This relation can also be found in the data. Short term job, i.e. those that ended within half a year, had payed initial wages which were below average by about 10.5%.

On the other hand, initial wages in medium- and long-term jobs were above average by roughly 13%. Clearly, the survival probability of a match is positively related to the initial wage that was payed. As mentioned previously, according to our model, this relation originates in agents' beliefs about the match quality which is a good predictor of the likelihood of survival. This in turn determines the value of a job and thereby affects the remuneration payed to the worker.

The lower panel in Table (6) reports summary statistics for different age cohorts. The first row contains the mean duration of an employment spell across age cohorts. What we observe is basically a mirror image of the first row in the upper panel. In short: the average duration of matches formed by young cohorts is substantially shorter than the jobs created by experienced workers. The average job duration more than doubles as we move from workers aged $18-20$ years to workers aged $34-38$ years. Not surprisingly, the mean value of ω is zero. This is by construction as we measure ω as the deviation of an individual's wage from the **average** wage in the **same** age cohort.

Arguably more interesting is the last row, i.e. the standard deviation of ω. We see a substantial increase in the standard deviation of ω implying that there is more dispersion in initial wages for experienced than for young workers. The dispersion roughly doubles as we move from the lower to the top age interval. This relation also holds when we use individuals' labor market experience instead of their age as the criterion. The observed increase in the dispersion can be explained by the stochastic nature of labor market transitions. Within a group of initially similar workers the degree of heterogeneity naturally rises over time as individuals all build up their own labor market histories. More heterogeneity in individuals' characteristics will be reflected by more dispersed initial wages.

7 Estimation Results

Coefficient estimates of the auxiliary model

In this section we report the coefficient estimates of the auxiliary model using observed data. Before we bring the model to the data we need to specify a grid for the job durations τ_n and age categories a_k. Clearly, the distribution of job separations across duration is strongly skewed to the left as most separations occur early in a firm/worker relationship. To account for that we choose a finer grid for short durations and coarser one for long durations. Choosing a grid that is too fine, however, has the disadvantage that relatively few observations fall in each interval leading to imprecise estimates of β.

Having in mind these considerations we specify the following grid $\{0, 1, 2, 3, 5, 7, 10\}$[19]. The choice of the age categories follows a similar logic[20]. As documented by Section 2, most of the labor market transitions are occur in the early years of an individual's career. To account for that, we make the age categories narrower for young individuals and wider for more experienced workers. In particular we choose the following grid for age $\{18, 20, 23, 26, 30, 34, 38, \}$. The coefficients to be estimated are

$$\beta = \begin{pmatrix} \eta_{18|20} & \eta_{20|23} & \eta_{23|26} & \eta_{26|30} & \eta_{30|34} & \eta_{34|38} & \eta_{\omega}, \\ \lambda_{0|1} & \lambda_{1|2} & \lambda_{2|3} & \lambda_{3|5} & \lambda_{5|7} & \lambda_{7|10} & \end{pmatrix}$$

The maximum likelihood estimator of β is given by the solution to the maximization problem stated in (12). Table (7) reports the results. Asymptotic standard errors are reported in parentheses.

Coefficient	Estimate	A.H. (%)	Coefficient	Estimate	M.E. (%)
$\lambda_{0\|1}$	0.4925 (0.0133)	0.885	$\eta_{18\|20}$	0.7992 (0.0345)	122.4
$\lambda_{1\|2}$	0.3187 (0.0095)	0.506	$\eta_{20\|23}$	0.6997 (0.0306)	101.3
$\lambda_{2\|3}$	0.2249 (0.0077)	0.339	$\eta_{23\|26}$	0.6089 (0.0308)	83.85
$\lambda_{3\|5}$	0.1638 (0.0058)	0.242	$\eta_{26\|30}$	0.4653 (0.0320)	59.24
$\lambda_{5\|7}$	0.1090 (0.0050)	0.153	$\eta_{30\|34}$	0.2664 (0.0336)	30.53
$\lambda_{7\|10}$	0.0876 (0.0044)	0.122	$\eta_{34\|38}$	0.0814 (0.0384)	8.480
			η_{ω}	-0.498 (0.0178)	-0.497

$\lambda_{\tau|\tau'}$ denotes the estimate of the baseline hazard rate for the interval $[\tau, \tau')$, $\eta_{a|a'}$ is the coefficient estimate for the age cohort $[a, a')$, η_{ω} is the coefficient estimate for the variable ω, A.H. denotes the average effect and M.E. is the marginal effect. Standard errors are in parentheses.

Table 7: Estimation results: Auxiliary model using observed data

The coefficients of the baseline hazard rates λ_i are all very precisely estimated and the values are in line with what one would expect. The probability of separating from an employer is decreasing in tenure. This is a standard prediction of models in which firms

[19]Job duration is measured in years.

[20]Recall that age in this framework refers to the worker's age at the time of the match formation and not at the time of separation.

and workers learn about the match quality. It is interesting to note though that the decrease is not linear in tenure. At the beginning of an employment relationship the baseline hazard rate is fairly high, but conditional on staying in the job it falls rather rapidly and subsequently levels out at a rate that is substantially lower than the initial value. This pattern is intuitive since most of the learning about the match quality takes place within the first few periods of an employment relationship. Matches that are revealed to be good continue to exist whereas bad matches break up. Thus, after an initial period of learning and selection each surviving match is less likely to be of bad quality which explains the non-linear decline in the baseline hazard rate. I also report the average hazard rate for each job duration. It is computed as the mean hazard rate of all observed individuals using the estimated coefficients. For instance the mean hazard in the population of separating from an employer in the tenure interval $[\tau_{i-1}, \tau_i)$ (conditional on still being in the job) is given by $\frac{1}{J}\lambda(i)\sum_{j=1}^{J} exp(\eta^j + \eta_\omega \omega_j)$.

Turning now to the estimates of the age coefficients η. Recall that each of the coefficients is linked to a dummy variable that is unity for the age interval that corresponds to the individual's age at match formation, and zero for all other age intervals. Recall furthermore that the "missing" age cohort, which serves as the means of comparison is given by the cohort with age ≥ 38 years. Therefore, any of the age coefficients reported in Table (7) has to be interpreted as the difference in the hazard between age groups k and K + 1, everything else equal.

In addition to the coefficient estimates I also compute the marginal age effects in the population. The results of which are reported in the column on the far right of Table (7). Both, the estimates and the marginal effects suggest that the age at match formation is a fundamental determinant of an individual's probability of separating from an employer.

A worker that starts a new job at the age of $18 - 20$ is more than twice (122.39% to be precise) as likely to separate from the current employer than a worker that starts a new job at (or above) the age of 38. This effect, however, declines dramatically with age, and especially within the first 10 years after labor market entry. An individual that starts a new job at the age of 30 is already about 60% less likely to separate from the current employer than it was 10 years before. For ages higher than 34 years the effect seems to vanish. The marginal effect of age in the cohort $34 - 38$ years is just 8.48%, implying that a job that is started at this age is almost as likely to end than a job that is started at age 38 years or higher.

According to the structural model, this effect on the separation probability is due to imperfect information in the hiring process. These imperfections are linked to an individual's labor market experience since for young - as opposed to experienced - workers there's less information available that facilitates the assessment of the workers' productive ability. Over time, as an individual stays in the labor market, more and more relevant information accumulates and hence the effect on the separation probability becomes less and less important. In the data this is reflected by a decreasing age coefficient. Moreover, the pattern in Table (7) suggests that the process which leads to a reduction in the importance of the age effect is strongly concentrated in the first 10 years after an individual's entry into the labor market. At later stages of the career, and in particular from age 34 years onward, it becomes rather negligible.

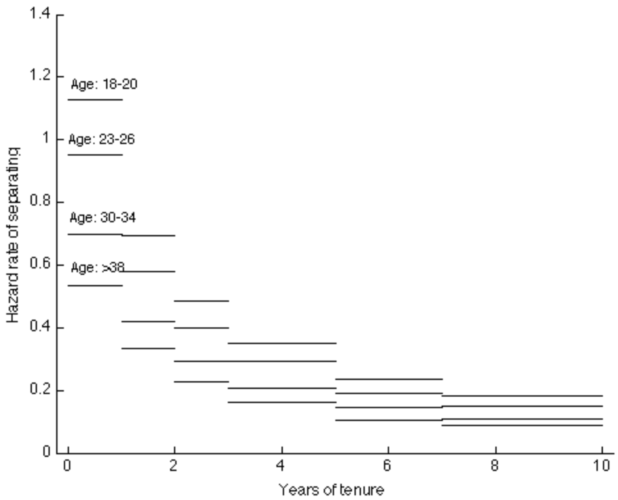


Figure 3

Actual hazard rate of separating from a job for selected age and tenure groups. Source: Own calculations based on data from the $NLSY$ 79.

The lower panel of Table (7) depicts the estimate of the coefficient of ω. The estimate is highly significant and has a negative sign. I also report the marginal effect which is computed as the mean percentage change in the hazard rate in the population induced by a +1% gap of individual's initial wage to the mean initial wage of the respective age cohort. The value of the point estimate implies that a positive 1% deviation of a worker's initial wage from the mean initial wage in the same age cohort leads to a reduction in the individual hazard by roughly 0.5%. Regarding the sign of η_ω, the structural model yields the same prediction.

This is based on the following mechanism: When forming a match agents observe a noisy signal about the unobserved quality of the match. This signal is used to draw inference about the true quality, resulting in a certain belief. Agents are rational in the sense that their beliefs are, on average, consistent with the true match quality. Good matches persist, bad matches break up. Consequently, the value of the initial beliefs is indicative for the probability of separating from an employer. For this reason they also determine the joint surplus of a job and thereby also the wage that is payed to the worker.

The hazard rates in Table (7) are computed across all age cohorts. Instead, Figure (3) reports the hazard rates for selected age cohort separately[21]. Two things are worth noting here: (1) At low job durations workers of all ages feature a high probability of separating from the employer. Though the probability is substantially higher for young workers than for more experienced ones. (2) Throughout all tenure classes the hazard rate for young workers is approximately twice as high as that for experienced workers. Therefore, also the percentage decline with tenure is roughly the same across cohorts, namely, around -84%. The relative difference is constant, the absolute difference in the hazard rates between young and experienced workers, however, shrinks dramatically from 0.5920 for newly created jobs to 0.0967 for jobs with $7-10$ years of tenure. Thus, it is clearly the case that newly created jobs are much more likely to break up when a young worker is involved. Interestingly, for established jobs, on the other hand, it matters much less if a young or an experienced worker is involved as the gap in the separation probability narrows down by quite a lot.

In order to assess the aptitude of the auxiliary model for capturing the main statistical properties of the data I perform a series of hypotheses tests. Checking the standard errors of the individual coefficient estimates can be considered as a first test to evaluate the explanatory power of the model. It appears that each of the estimates is statistically different from zero at the 98% confidence level[22].

First, I perform a likelihood ratio test of the hypothesis that the age coefficients are all the same, i.e. $\eta_{18|20} = \eta_{20|23} = ... = \eta_{34|38}$. The value is 446.8 with 5 degrees of freedom which rejects the hypotheses at the 99% confidence level. At the same level of confidence I reject the hypotheses that the baseline hazard rates are all constant across tenure intervals, i.e. $\lambda_{0|1} = \lambda_{1|2} =, ..., = \lambda_{7|10}$. Next, I test for the hypotheses that the

[21] Table (8) in Appendix D contains the corresponding numbers for all the age groups we consider in the estimation.

[22] In a preliminary step I also included a worker's total labor market experience (measured as the worker's cumulated number of weeks being employed since labor market entry) as an explanatory variable. This measure, however, turned out to be highly correlated with the individual's age variable and thus causing problems of multicollinearity.

hazard rates for two consecutive intervals are the same, i.e. $\lambda_{\tau|\tau'} = \lambda_{\tau'|\tau''}$. The hypotheses $\lambda_{0|1} = \lambda_{1|2}$, $\lambda_{1|2} = \lambda_{2|3}$, $\lambda_{2|3} = \lambda_{3|5}$, $\lambda_{3|5} = \lambda_{5|7}$ are, respectively, rejected at the 99% level whereas $\lambda_{5|7} = \lambda_{7|10}$ is rejected at the 97.5% level. Next, I perform a similar test to check whether the age coefficients of two consecutive age cohorts are constant, i.e. $\eta_{a|a'} = \eta_{a'|a''}$. There the test rejects $\eta_{18|20} = \eta_{20|23}$ at the 97.5% confidence level whereas the remaining hypotheses, $\eta_{20|23} = \eta_{23|26}$, $\eta_{23|26} = \eta_{26|30}$, $\eta_{26|30} = \eta_{30|34}$, $\eta_{30|34} = \eta_{34|38}$, are all rejected at the 99% level.

Lastly, I test whether the choice of the age grid was an appropriate one. To this end, I first estimate the auxiliary model using a finer grid. Then I use the associated value of the likelihood function to perform a likelihood ratio test of the hypotheses that the grid in the original specification is as good as the finer one. Here, the term "good" refers to the difference in the likelihood of both specifications. When using the grid $\{18, 19, 20, 21, 23, 26, 28, 30, 32, 34, 38\}$ I can reject the hypotheses that the original grid - $\{18, 20, 23, 26, 30, 34, 38, \}$ would perform as good at the 75% confidence level only. Extending the grid to include also higher ages, i.e. $\{18, 19, ..., 24, 26, ..., 44\}$, leads to the same result.

The results of all those tests make me confident that the chosen auxiliary model provides a good statistical description of the underlying data along the dimensions that are needed for the identification of the model's structural parameters. Those dimensions are meant to capture any age-cohort effects that - according to the model - are causal for differences in the job separation and job turnover behavior of (a) individuals of different age and (b) along the working life cycle of each individual.

Estimating the parameters of the structural model

Estimating the structural parameters means finding a $\tilde{\phi}_S^H$ that solves the minimization problem stated in (13). Performing one evaluation of the objective function involves the following steps:

1. Given the structural parameters ϕ the model is solved for the optimal policy functions.

2. With those at hand we can simulate the life path of a number of individuals to obtain data on the individuals' employment spells.

3. Using the simulated data we then estimate the parameters of the auxiliary model and finally evaluate the quadratic form in (13).

In order to solve the model we need to take a stand on a variety of issues. Concerning an individual's life-cycle, we need to specify (a) the survival probabilities ρ_k, for $k = \{1, 2, ..., \bar{K}\}$ and (b) the number of periods an individual stays in the labor market, set $\bar{K}$. The data that is used here does not allow for the identification of individuals' death. We do observe individuals dropping out of the survey, however, the cause of dropping out is not specified. It is unlikely though that death plays an important role. The individuals observed in 2006 (the year of the last survey) are aged $42 - 48$ years. According to the U.S. National Center for Health Statistics those individuals can expect to have $36.9 - 30.5$ more years to live. Moreover, the annualized probability of death for this age cohort is less than 0.25% indicating that, for those individuals, the risk of death is fairly low. In light of that, we set $\rho_k = 1$ for $k = 1, ...\bar{K}$. Furthermore, the time dimension of the $NLSY79$ survey is still limited in the sense that we do not observe individuals up until retirement. Consequently, we can not compute $\bar{K}$ from the survey data. Instead we set $\bar{K} = 45$, implying that individuals spend at maximum 45 years in the labor force.

In the simulation step we simulate the working life cycle of $\mathcal{S} = 10^5$ individuals. For each individual we observe a certain number of employment spells. Clearly, the number and the length of which depend on the set of parameters ϕ[23]. The number of observations in the estimation is therefore endogenous and depending on the set of parameters. This is undesirable since a low number of observations implies relatively inaccurate estimates of the structural parameters. To circumvent this problem I introduce a penalization term into the algorithm. In particular, the algorithm penalizes those parameter constellations for which the implied number of observations is lower than a certain critical value. As a critical value I use $2 * \mathcal{S}$.

Furthermore, the number of worker types is assumed to be two. The types are denoted by $i \in \{i_1, i_2\}$. Before we bring the structural model to the data we need to make the following adjustment. Notice that the variance of the noise component is the only characteristic that distinguishes young from experienced workers. Therefore, the model would attribute all the differences in the job separation across age entirely to the information imperfections. This is clearly an overstatement of the importance of the information channel. Estimates based on the current specification would most likely be biased. Obviously, there are also other sources that cause separation rates to differ across age. To account for that we do not ex-ante restrict (a) the exogenous rate of separation σ and (b) the job quality, governed by α, to be the same across types. Instead, we allow σ and

[23]If, for instance, the policy functions implied by ϕ lead high threshold beliefs than few individuals will accept job offers and the number of employment spells in the resulting data set will be relatively low.

α to vary across i-types, i.e. we consider $(\sigma_{i_1}, \sigma_{i_2})$, $(\alpha_{i_1}, \alpha_{i_2})$ and estimate them separately.

Moreover, I restrict the process that governs changes in the workers' type to be irreversible. In particular, I set $\mu_{2|2} = 1$, $\mu_{2|1} = 0$, which excludes the possibility of type "depreciation". This assumption does not seem to be very restrictive. A worker's type is meant to capture factors that affect the screenability of a worker such as the stock of information about the worker's labor market history. The more information is around, the better an applicant can be screened and categorized by the hiring firm. Such information hardly gets lost over time. Therefore, it is arguably a minor point to neglect type depreciation. The only free parameter that governs changes in the workers' type is therefore $\mu_{1|2}$. Notice that $\mu_{1|1}$ can be recovered using $\mu_{1|1} = 1 - \mu_{1|2}$ once we know $\mu_{1|2}$.

In total the model consists of 14 structural parameters. Those are:

$$\left(\beta, \tilde{\lambda}, e, y^g, y^b, b, \mu_{1|2}, \varphi, \sigma_{i_1}, \sigma_{i_2}, \alpha_{i_1}, \alpha_{i_2}, s_{i_1}, s_{i_2}\right)$$

Ten of which are estimated. Those are $\left(y^b, b, \mu_{1|2}, \varphi, \sigma_{i_1}, \sigma_{i_2}, \alpha_{i_1}, \alpha_{i_2}, s_{i_1}, s_{i_2}\right)$. The remaining four, i.e. $\left(\beta, \tilde{\lambda}, e, y^g\right)$, are set beforehand. We set $y^g = 1$ which is basically a normalization. What matters for the expected per-period output of a match is not the level of y^g but the difference $y^g - y^b$. Therefore, we fix y^g an estimate y^b. We set the personal discount factor β equal 0.9569 which implies an annual interest rate of 4.5%. It is typically hard to get reasonable estimates for the personal discount factor β. Especially in a setting like ours the existence of individuals' heterogeneity and a finite life cycle renders it hard to estimate β within meaningful bounds. We set the firm's bargaining power e equal to 0.5 which is a standard value used in the search and matching literature. The estimation strategy used here is unlikely to properly identify the arrival rate of new job offers $\tilde{\lambda}$. This parameter is a key determinant of the duration of an unemployment spell but it has little effect on the duration of an employment spell. The hazard rate approach which we adopted focuses on the duration of jobs but it does not process any information on individuals' unemployment spells. Consequently, $\tilde{\lambda}$ would be poorly identified and any estimate of it would be meaningless. We set $\tilde{\lambda} = 1$ which we believe is a reasonable choice given that we use yearly observations and the average unemployment spell in the *U.S.* is significantly shorter than a year.

To estimate the parameter vector $\tilde{\phi}_S^H = \left(y^b, b, \mu_{1|2}, \varphi, \sigma_{i_1}, \sigma_{i_2}, \alpha_{i_1}, \alpha_{i_2}, s_{i_1}, s_{i_2}\right)$ I use a hybrid method that was proposed by Nagypal (2007). This method first evaluates the objective function along a relatively coarse grid. The initial grid that I chose has dimension $4 \times 2 \times 6 \times 4 \times 3 \times 3 \times 4 \times 4 \times 3 \times 3$. To economize on computational time I start out with a low number of simulations. In each subsequent round I increase the number of simulations leading to more precise evaluations of the objective function. However I consider just those grid points that perform well in the previous round. All the remaining points are eliminated. After this refinement I am left with a grid with 279 points on the grid. The values of the objective function on this grid range from 6.46×10^{-6} to 2.15×10^{-5}. The method is described in detail in Appendix E of Nagypal (2007)[24]. In the next step I use the remaining grid points as starting values in a simplex algorithm that solves the minimization problem in (13). Table (8) reports the estimates of the structural parameters.

Parameter	Explanation	Estimate	
y^b	Output of a bad firm/worker match	-3.2433 (0.3350)	
b	Flow value of unemployment	0.4333 (0.1233)	
$\mu_{1\|2}$	Probability of a type change	0.4134 (0.0726)	
φ	Probability that the match quality is revealed	0.6157 (0.0792)	
Parameter	**Explanation**	**Estimate**	
		Type 1	Type 2
σ_i	Probability of exogenous job separation	0.2497 (0.0122)	0.1084 (0.0323)
α_i	Mean of agents' beliefs	0.3781 (0.0969)	0.7571 (0.0706)
s_i	Standard deviation of agents' beliefs	0.0962 (0.0097)	0.2030 (0.0156)

Estimates of the structural parameters of the model presented in Section 3. The standard errors are reported in parentheses.

Table 8: Estimation results: Structural Model

To test whether the structural model is well specified I perform a global specification test that was proposed by Gourieroux et al. (1993). This test is based on the minimized value of the objective function with the statistic

[24]Using the terminology by Nagypal (2007) I use a total of 4 rounds in which the number of simulations is $\{5 \times 10^3, 10^4, 5 \times 10^4, 10^5\}$.

$$\zeta_S = \frac{SH}{1+H} \times \min_{\phi} \left(\hat{\beta}_J - \frac{1}{H} \sum_{h=1}^{H} \beta^h(\mathbf{y}_S, \phi) \right) \mathcal{Q}^* \left(\hat{\beta}_J - \frac{1}{H} \sum_{h=1}^{H} \beta^h(\mathbf{y}_S, \phi) \right)' \tag{16}$$

that it is asymptotically distributed as a χ^2 with $dim(\beta) - dim(\phi)$ degrees of freedom. The test statistic in our case is equal to 1.374 which is well below the critical value of 7.82 for the 95% confidence level with 3 degrees of freedom. Therefore, we can not reject the null hypotheses that the model is well specified. The standard errors depicted in Table 8 are computed by evaluating the expression in (14). For that we consider a 0.0001% deviation from the estimates in Table 8. To get accurately estimated standard errors we rely on a very large set of simulated data. In particular we simulate the working life cycle of 20 million individuals which results in around 120 million observations of employment spells. The standard errors in Table 8 suggest that the coefficients are estimated fairly precisely. This does not, per se, rule out the possibility that the minimum that was found is local minimum.

To double-check the accuracy of the estimation procedure and in particular to identify flat regions of the objective function surface we compute the objective function fixing all but one parameter at the estimated value and varying the remaining parameter over a range of feasible values. The results are depicted in Figure 8 in Appendix D. Most importantly, we do not find any notable flat regions of the criterion function which confirms the accuracy of our estimates[25]. The possibility of being of being trapped in a local minimum is pretty much ruled out by the operating mode of the grid search algorithm which we use in the first place to obtain the starting values. This algorithm provides us with a range of values, each of which is potentially close to a (local) minimum. Using many different starting values in the subsequent simplex-search algorithm guarantees that we

[25] An interesting case emerges when we vary s_2. The panel in the last row in Figure 8 illustrates that for values of $s_2 > 0.26$ the criterion function soars up. This is due to the following reason. s_2 is defined as the standard deviation of a truncated normal distribution with domain $[0, 1]$. To draw random numbers that follow a truncated normal distribution we have to compute the standard deviation (together with the mean) of the corresponding un-truncated normal. The issue is, however, that for a truncated normal with mean in $[0, 1]$ there exists an upper bound on the standard deviation which is around 0.288. Any arbitrarily high value for the standard deviation of the un-truncated normal translates into a value for s that is always below this upper bound. Consequently, for values of s_2 close to that upper bound the implied random draws for agents' beliefs are that much dispersed so that, in the end, very few draws fall above the threshold belief. This means that very few individuals get matched and in the resulting simulated data we observe very few employment spells. As mentioned previously, we incorporated a penalty term into the algorithm that adds a very high positive number to the objective function whenever the number of simulated employment spells falls short of a critical value. The case of s_2 being close to the upper bound is such a case, hence we observe the criterion function jumping upwards.

can identify (and dismiss) local minima by comparing the values of the criterion function at all the points at which the algorithm terminates the search process.

The estimate of $\mu_{1|2}$ is equal to 0.4134, meaning that an employed type 1 worker faces a 41.34% (annualized) probability of becoming a type 2, conditional on staying employed. Consequently, a young worker would need, on average, 2.42 years of tenure on a single job before receiving the type-upgrade. However, jobs often do not survive that long and employment spells are typically interrupted by periods of unemployment. Especially the high turnover for young cohorts is detrimental to a quick upgrade as it prevents the cumulation of years of tenure on a given job. Thus, it takes substantially longer than 2.42 years until a type change occurs. In the simulated version of the model (an explanation follows), where we take into account all the labor market transitions of an individual, we find that the average worker moves from type 1 to type 2, 5.25 years after labor market entry. The quality of a match is initially unknown but firm learns about it, over time, on the basis of observed output. Each period there is a constant probability that the match quality is fully revealed. This probability is estimated to be $\varphi = 0.6157$. As a result, it takes, on average, 1.62 years until the firm learns the true quality of the match.

The lower panel of Table (8) reports the estimates of the parameters that we allowed to differ across i types. In our setup, σ is the rate at which matches break up for exogenous reasons. In the estimation it can be thought of a residual that captures all the separations that are not due bad match quality. The (annualized) rate of exogenous match separation for matches with type 1 and type 2 workers is estimated to be, respectively, 0.2497 and 0.1084. The value for experienced workers is close to what empirical studies usually report. Davis and Haltiwanger (1992), for instance, find that in the U.S. the rate of exogenous job destruction is 11.3% per year. The rate for young cohorts, that we find, is substantially higher than the rate for experienced worker. These numbers imply that matches of experienced workers are hit, on average, after 9.2 years by an exogenous shock leading to a break-up, whereas matches of young workers are hit after about 4 years. To test whether the estimates of σ_i are actually statistically different from one another I perform a Wald test of the hypotheses that $\sigma_1 = \sigma_2$, i.e. the rate of exogenous separation is the same across types. The statistic associated the with null hypotheses is given by

$$\zeta^W = c(\tilde{\phi}_S^H)\,(C\mathcal{W}C')^{-1}\,c(\tilde{\phi}_S^H)$$

where $\zeta^W \sim^a \chi_2^2$ and $\mathcal{W}$ is the asymptotic variance-covariance matrix determined in (14). $c : \Phi \to \Re$ is a continuously differentiable function on the parameter space Φ that incorporates the null hypotheses: $H_0 : c(\phi_0) = 0$ and $C(\tilde{\phi}_S^H)$ is the gradient of $c(\cdot)$

evaluated at the optimal unconstrained estimator $\tilde{\phi}_S^H$. The test statistics associated with the null hypotheses $\sigma_1 = \sigma_2$ is 10.08 which is above the critical value of 9.2 for the 99% confidence level with 2 degrees of freedom. Therefore, we reject the null hypotheses that the exogenous rate of job separation is the same for type 1 and type 2 workers.

In our model α has a threefold interpretation. On the one hand, it is the mean of the distribution from which the true - but unknown - job quality π is drawn from. Given the zero-mean assumption for the noise component ϵ, is also represents the mean of the observed signal γ, and provided how agents form their beliefs about π it is, at the same time, also the mean of the beliefs that are realized. The estimates we obtain for α_1 and α_2 imply that the true job quality is substantially lower for young as compared to experienced workers. Notice, however, that α is not the average quality of all *realized* but of all *potential* matches. Not all matches are formed in equilibrium. Encounters with a signal below the threshold do not result in the match formation. Hence the actual average quality of existing match is higher, and in our case it is equal to 0.524 for type 1 workers and 0.894 for type 2 workers. This says that almost half of all employed type 1 workers but just 10% of type 2 workers are in a mismatch[26].

The large difference between the two groups can be explained by the type-specific noise component. Recall that young workers emit a less informative signal than experienced worker. A poor signal that is very noisy is more likely to indicate a good match than the same signal that is precise. Therefore, firms can risk to hire young workers also with relatively poor signals as there is substantial likelihood that the resulting matches are in fact good. Consequently, the hiring threshold is lower for type 1 workers and so is the average quality of existing matches. Lastly, I perform a Wald test on the hypotheses $H_0 : \alpha_1 = \alpha_2$ to check whether the difference in the estimated values is statistically significant. The test statistic is 13.09, hence we reject the hypotheses that the job quality is the same across worker types at the 99% level of confidence.

The goal of this paper is to assess the importance of worker-specific informational frictions for explaining the observed differences in labor market outcomes across age cohorts. In our model those frictions are captured by a noise component that distorts the workers' signal. The worker specificity shows up in the variance of the noise component $\sigma_{\epsilon,i}^2$, which we assume to be different across worker types. $\sigma_{\epsilon,i}^2$ determines the variance of agents' beliefs via $s_i^2 = \sigma_\pi^4 / \left(\sigma_\pi^2 + \sigma_{\epsilon,i}^2\right)$. We did not estimate $\sigma_{\epsilon,i}^2$ itself but we can back it out from the expression $s_i = \sigma_\pi^2 / \sqrt{\sigma_\pi^2 + \sigma_{\epsilon,i}^2}$ using the estimated values for the standard deviation

[26]Notice that here we are referring to matches whose quality was not yet revealed.

of agents' beliefs s_i. The estimates of s_1 and s_2 are respectively 0.0962 and 0.2030. We use a range of values for σ_π^2 and find that the ratio $\sigma_{\epsilon,1}/\sigma_{\epsilon,2}$ takes values of around 2.1. The implied signal to noise ratio is about 4.5 times higher for experienced workers as for young individuals. Lastly, I perform a Wald test to check whether the difference in the precision of workers' signals is significant. We find that the test statistics associated with $H_0 : s_1 = s_2$ is 12.78, thus we reject the hypotheses that the informational friction is **not** worker-type-specific.

Discussion of the Results

In this section we assess how well the predictions obtained from the structural model match up with actual data. To that end we use the estimated parameters and simulate the model to generate data on the individuals' working life cycle. From the simulated data we compute a variety of statistics which we can then compare to their counterparts in the observed data. Before doing that we take a glance at the optimal policy functions implied by the model. Recall that the policy in this model is given by the threshold value of beliefs for both worker types. Figure (4) depicts the results.

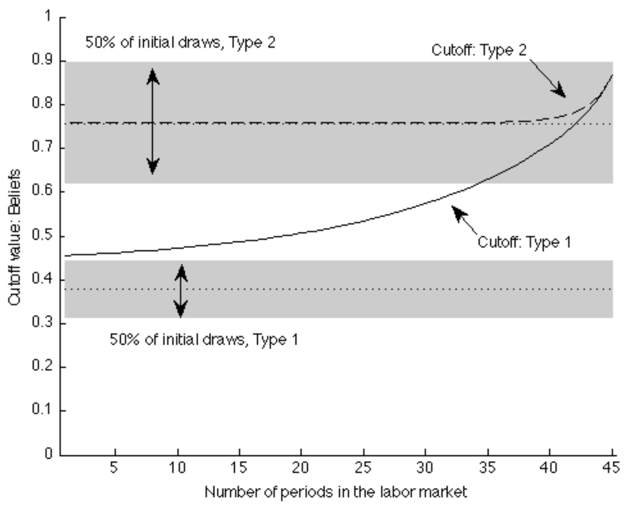


Figure 4

(a) Equilibrium threshold value of beliefs for type 1 and type 2 workers (solid and dashed line), (b) mean value of the initial beliefs (dotted lines), (c) 50% confidence bounds around the mean (shaded area).

As expected, the cutoff for both types, is increasing in age. This is intuitive. An individual getting ever closer to the exit age becomes less and less valuable for a (new) match as the time horizon, over which a firm can collect the surplus, shrinks. To compensate firms for the shorter horizon - and for the risk that the match might be bad - a higher surplus is required to make the matching with older workers attractive. Consequently, the threshold increases. Notice, however, that the threshold is flat for almost the entire career of type 2 workers. For type 2 workers there is very little risk that the match turns out to be bad, hence firms do not require any compensation.

The threshold for type 2 is higher for type 1. The intuition is as follows. The outside option of employed type 2 workers is more valuable as their expected value of a job is higher than that of a type 1 worker. Therefore, they can extract a higher fraction of the total surplus which translates into an increased reservation belief. Figure 4 also contains the mean value of the initial beliefs (indicated by the dotted line) and the 50% "confidence bounds" around the mean (shaded area).

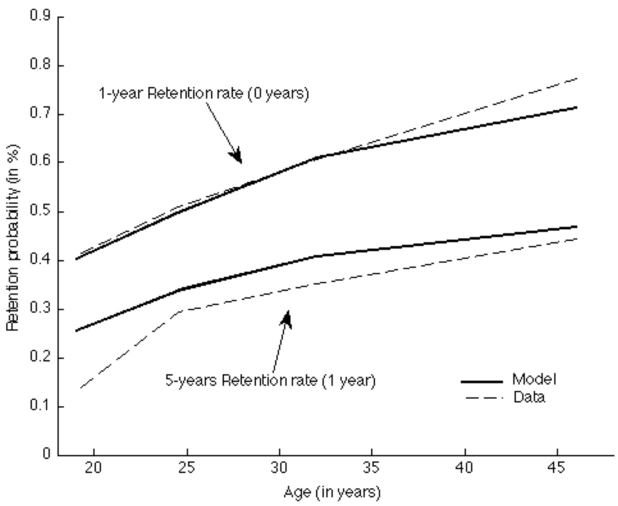


Figure 5

The 1-year retention rate for newly employed workers (upper two lines) and the 5-years retention rate for workers with one year of tenure (lower two lines). Comparing the results obtained from the model (solid line) with the counterpart in the data (dashed line). Source: Own calculations using data from the $NLSY$ 79.

These bounds indicate the region in which 50% of all realized initial draws fall into. From there, we clearly see that the initial beliefs of type 1 workers are substantially less dispersed. This is a consequence of how beliefs are formed (see Equation (2)). More noise in the signal means that - when forming beliefs - more weight is put on the mean value and less on the actual signal. Thus, there is less dispersion in type 1 beliefs. Notice furthermore that the cutoff value for type 1 is above the shaded area, meaning that type 1 workers are, on average, less successful in generating a signal that sufficient for a match formation. Thus the noise in the signal (combined with a lower mean value of beliefs α_1)

is basically a hiring barrier for those workers[27].

One of the main findings in Section 2 was that young individuals are substantially more likely to separate from a new job than experienced workers. We computed job retention rates from the data and found that the likelihood of leaving a new job declines with age. Now we compute the same statistics using the simulated data. Figure (5) reports the results for: the 1-year retention rate for newly employed workers (the upper two lines) and the 5-years retention rate for workers with one year of tenure (lower two lines)[28]. The solid line is the prediction of the model whereas the dashed line is the counterpart in the data.

Duration / Age	[18-20)	[23-26)	[30-34)	[38-
1 year	0.41 **(0.40)**	0.51 **(0.49)**	0.60 **(0.61)**	0.77 **(0.71)**
2 years	0.21 **(0.25)**	0.33 **(0.36)**	0.42 **(0.48)**	0.61 **(0.59)**
3 years	0.12 **(0.20)**	0.26 **(0.31)**	0.33 **(0.41)**	0.45 **(0.51)**
4 years	0.09 **(0.17)**	0.21 **(0.26)**	0.27 **(0.36)**	0.41 **(0.45)**
5 years	0.07 **(0.15)**	0.17 **(0.23)**	0.24 **(0.32)**	0.36 **(0.41)**

Job retention rates for newly employed workers: Rows represent different retention periods, columns indicate workers' age at the time of recruitment. Numbers in bold (normal font) represent the model (data).

Table 9

Table (9) contains the results for the remaining retention classes. For all the cases, the model can capture very well the increase the in the retention probability with age. Also the level that is predicted is broadly in line with the actual data, though for young workers, the model slightly overpredicts the short-horizon retention rate. For experienced workers the predictions of the model are generally closer to the empirical facts.

We also compute the retention probabilities for workers with one year of tenure. Table (14) in Appendix C reports those results. The pattern for that tenure class is generally very similar to that of newly employed workers.

[27]This has important implication for the duration of unemployment. According to the model, we should observe young (type 1) workers having longer spells of unemployment than experienced workers. In reality, this effect, however, might - at least partly - be offset by a higher arrival rate of job offers for young as they are typically less specialized yet and hence more willing to take on any given job. In our model we abstract from this issue thus we can not account for age-specific differences in the duration of unemployment.

[28]Recall that the x-years retention rate is the probability that a worker with a certain tenure keeps her current job for x more years.

An aggregate implication of unstable jobs is high job turnover among young cohorts. In the data this is clearly observable as illustrated in Section 2. In what follows we presents some tables and figures documenting how well our model can account for observed differences in job turnover across age cohorts. In a world with high turnover we should observe (a) short short employment spells and (b) individuals holding many jobs in a given period of time.

Tenure / Age	[18-20)	[23-26)	[30-34)	[38-
$0-2$	0.89 **(0.84)**	0.61 **(0.39)**	0.37 **(0.23)**	0.24 **(0.21)**
$2-5$	0.11 **(0.16)**	0.27 **(0.39)**	0.27 **(0.26)**	0.21 **(0.22)**
$5-10$	0.00 **(0.00)**	0.12 **(0.22)**	0.24 **(0.31)**	0.22 **(0.24)**
> 10	0.00 **(0.00)**	0.00 **(0.00)**	0.12 **(0.20)**	0.33 **(0.33)**

Proportion of workers with $\tau \in \{[0,2),[2,5),[5,10),>10\}$ years of tenure within a given age cohort. Numbers in bold (normal font) represent the model (data).

Table 10

Table (10) reports - for different age cohorts - the fraction of workers within a given tenure class. The numbers in bold represent the model outcome. Clearly, young workers are over-represented in short-term jobs as, for them, little time has passed since entry into the labor market. Then, as the career progresses, more and more workers move into medium- and long-term jobs. Actually, for the youngest cohort the model can replicate fairly well the actual tenure "distribution". It over-predicts the fraction of young workers that move into medium-term jobs (second column). Also the results for more experienced workers are broadly in line with the empirical facts.

Age	[18-20)	[20-23)	[23-26)	[26-30)	[30-34)	[34-38)	[38-
				Data			
	0.332	0.484	0.629	0.746	0.806	0.882	0.963
	(0.211)	(0.236)	(0.232)	(0.211)	(0.201)	(0.165)	(0.084)
				Model			
	0.211	**0.376**	**0.539**	**0.661**	**0.765**	**0.854**	**0.948**
	(0.185)	(0.209)	(0.213)	(0.203)	(0.185)	(0.156)	(0.093)

Each entry represent the average number of jobs a worker holds until age a as a fraction of the total number of jobs hold in the entire career. The standard errors are in brackets.

Table 11

One of the key facts presented in Section 2 indicated that there is a lot of job-shopping among young but substantially less turnover for experienced workers. As a consequence, Individuals typically hold the vast majority of all their lifetime jobs within the first 10 years of the career. In Table (11) we report the average number of jobs a worker holds until a certain age as a fraction of the career total. The upper panel shows the data, the lower panel contains the numbers implied by the model.

Generally speaking, the model can replicate quite well the observed pattern, i.e. the steep increase in the number of jobs during the early years in the career and the flattening out in later stages. However, as mentioned previously, the model slightly underestimates the amount of turnover for young workers. Therefore, the initial increase is less pronounced than in the data. However, there is substantial heterogeneity in the number of jobs individuals hold along the career. This is indicated by the large standard deviation. So in order to get a better picture, Figure (6) plots the mean number of jobs together with the 50% confidence area around the mean. This area determines the region in which 50% of the individuals belonging to a given age cohort fall into. From there, we see that there is, actually, a lot of overlap between the model outcome and the actual data.

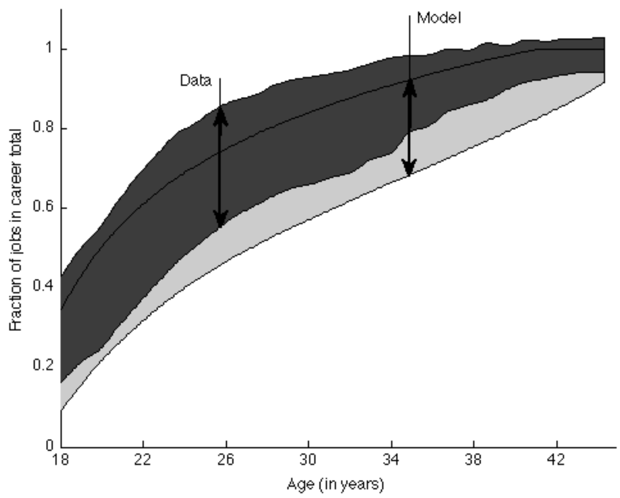


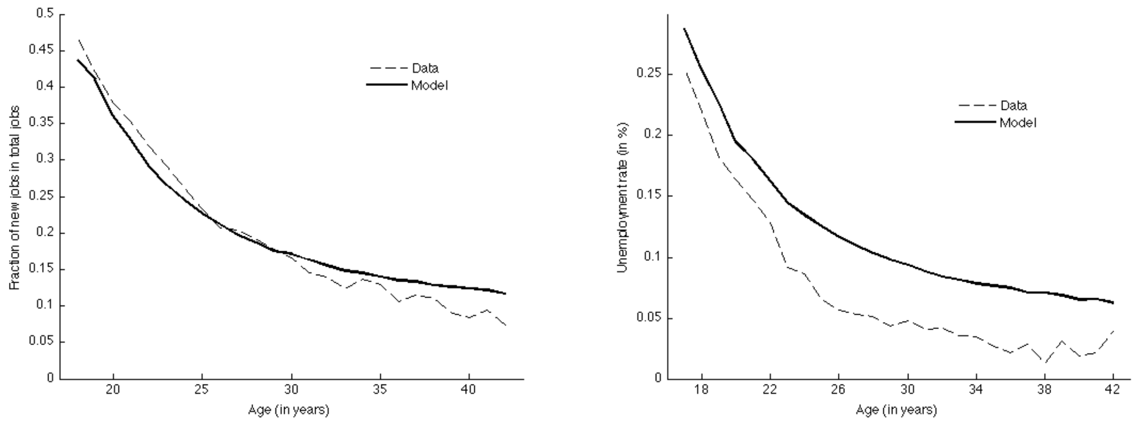
Figure 6

The 50% confidence area around the average number of jobs a worker holds until a certain age as a fraction of the career total. Comparing the model (light-grey area) with the data (dark-grey area). Source: Own calculations using data from the $NLSY$ 79.

Next, we report the number of newly employed workers as a fraction of total employment. Panel (a) in Figure (7) compares the model outcome with the empirical counterpart. In the data we define a newly employed as a worker having at most 20 weeks of tenure. In that dimension the model does extremely well. Table (10) has shown that it generally underestimates the proportion of workers with tenure less than 2 years but it is highly accurate predicting the very short run horizon. Lastly, we depict the age-specific unemployment rate, obtained from the model, and compare it to the data. Panel (b) in Figure 7 contains the respective plots.

Qualitatively, the model can capture the observed decline in unemployment with age. However quantitatively, the model overestimates unemployment for, basically, all age groups. Actually, this is not too surprising. The empirical strategy, that we have used to estimate the structural parameters of the model, uses information just on individuals' employment spells, but is does not make use of information about unemployment spells. Thus, there is no target in the auxiliary model that captures the observed pattern of unemployment.

In the previous section we found that the model account fairly well for the observed age-specific differences in labor market outcomes. We also showed that all the parameters we allowed to differ across i-types, i.e. σ, α and s, are actually statistically significant from each other. What remains is the question about the contribution of each of those parameters to explaining the observed life-cycle dynamics of individual job mobility.



(a) (b)

Figure 7

Panel (a): The number of newly employed workers with a given age as a fraction of total employment (of the same age group). Comparing the results obtained from the model (solid line) with the empirical counterpart (dashed line). Panel (b): Comparing the unemployment rate generated by the model (solid line) with the empirical counterpart (dashed line). Source: Own calculations using data from the $NLSY$ 79

To address this question we run an experiment in which we implement various restrictions on the parameters and compare the model outcomes obtained under the restrictions with (a) the data outcome and (b) the results of the unrestricted model. For that we take the results for the job retention probabilities as a standard of comparison. In particular, we use the 1, 3 and 5-years retention rates for newly employed workers. The results are reported in Table 12. Each of the the first five columns depicts the difference in the $\{1,3,5\}$-years retention rate between workers with a given age (top row) and workers aged $18-20$ years.

What we capture is basically the increase in job stability over the life cycle. The first row in each panel is the data outcome. There the first entry, for instance, means that newly employed workers aged $23-26$ years are 10.1% more likely to retain the job than newly employed workers aged $18-20$ years. Each of the rows thereunder contain a particular model outcome obtained under a given parameter restriction.

Age	[23-26)	[26-30)	[30-34)	[34-38)	[38-	% explained
			1-year job retention rate			
Data	**0.101**	**0.139**	**0.198**	**0.245**	**0.364**	
-	0.000	0.000	0.000	0.000	0.000	0.000
s	0.026	0.055	0.092	0.122	0.141	0.418
α, s	0.077	0.135	0.192	0.233	0.272	0.873
σ, α, s	0.091	0.146	0.205	0.261	0.313	0.985
			3-years rate			
Data	**0.141**	**0.164**	**0.220**	**0.272**	**0.329**	
-	0.000	0.000	0.000	0.000	0.000	0.000
s	0.018	0.046	0.069	0.095	0.113	0.287
α, s	0.058	0.108	0.152	0.186	0.221	0.624
σ, α, s	0.095	0.155	0.212	0.265	0.316	0.906
			5-years rate			
Data	**0.096**	**0.118**	**0.169**	**0.189**	**0.289**	
-	0.000	0.000	0.000	0.000	0.000	0.000
s	0.014	0.028	0.046	0.064	0.076	0.259
α, s	0.039	0.072	0.104	0.124	0.152	0.562
σ, α, s	0.078	0.127	0.172	0.212	0.261	0.992

Each entry in the first five columns represent the difference in the $\{1,3,5\}$-years retention rate between workers with a given age and workers aged $18-20$ years. The first row in each panel is the data. The following rows represent the model outcome obtained when restricting certain type-specific parameters to be the same across i-types. We report the parameters that are allowed to differ across i-types. The leftmost column reports the average agreement with the data-outcome.

Table 12

The restrictions we implement are as follows: (1) $\sigma_i = \overline{\sigma}, \alpha_i = \overline{\alpha}, s_i = \overline{s}$, $\forall i$ (second row). This means that we do not allow for any type-specificity of parameters. (2) $\sigma_i = \overline{\sigma}, \alpha_i = \overline{\alpha}$ (third row). In this case we allow just the noise component to be different across i types. (3) $\sigma_i = \overline{\sigma}$ (fourth row). Here we allow α and s to differ. The last row depicts the outcomes of the unrestricted model which we obtain by using the parameter estimates from Table 8. In each of the restrictions we fix the parameter at some value which is given by the arithmetic mean of the estimated, type specific, values. For instance $\overline{\alpha} = (\alpha_1 + \alpha_2)/2$. The leftmost column in Table 12 reports the average agreement of the respective scenario with the data-outcome.

For brevity we provide a description just of the results for the 1-year job retention rate, i.e. the first panel in Table 8. Notice that the unrestricted model (fifth row) can account, on average, for 98.5% of the increase in the observed 1-year rate. When we restrict all the parameters to be the same across i-types (second row) the model is not able to generate any life-cycle dynamics. The zeros indicate that, in this scenario, there is not change in the retention rate as workers age. Consequently, all of the age-specificity remains unexplained. In the next case (third row) we still keep σ and α fixed but we allow s to differ across i-types.[29].

Introducing worker-specific information imperfections helps considerably improving the empirical fit of the model. 41.8% of the observed increase in the 1-year retention probability can be accounted for only by the worker-specificity of the imperfections. The contribution is lower but still substantial for the 3-years and 5-years retention rates. Keeping σ still fixed but allowing, in addition to s, also the job quality α to differ, adds another 45.5%. The remaining 11.2% can be explained by introducing worker-type-specific rates of exogenous match separation.

Clearly, the major part of the observed increase can be explained by introducing worker-type-specificity in α and s. Both parameters are fundamental for the signal extraction process as α represents the mean of the observed signals and s relates to the amount of noise in a signal. The contribution of σ seems negligible, at least for the short-term horizon. However, when looking at the results depicted in the mid- and the lower panel of Table 12 it becomes clear that contribution of σ rises substantially when we consider the medium-and long-term horizon. For the 3-years and 5-years retention horizon the contribution of σ is, respectively 28.2% and 43%. There is an intuitive explanation for why σ is less important for explaining the short-run separation pattern but substantially more important for the medium- and long-run horizon.

[29] In all the cases in which we allow parameters to differ across types we set them equal to the respective estimated value.

In our model all the mismatches break up upon the revelation of the bad match quality. According to the estimates, the quality of a match is revealed, on average, after 1.62 years but an exogenous break-up occurs, on average, only after 4 years (young) and 9.2 years (experienced). Therefore, most of the match separation in the short run is due to endogenous separation whereas break-ups in the medium- and long-run are driven by exogenous shocks. Furthermore, what matters for the magnitude of endogenous break-ups is primarily the distribution from which the signals are drawn, and this distribution is shaped by α and s. As a result it is α and s that govern the separation process in the short run whereas σ drives job destruction the medium and long run.

8 Conclusion

In this paper we present new empirical evidence documenting the life-cycle dynamics of individual job mobility in the U.S. labor market. Based on *NLSY* 79 data we find that the job retention probability for newly employed workers is strongly increasing with an individual' age and labor market experience. We estimate that the first 10 years of labor market experience raise the probability of retaining a new job for one year or longer by roughly 20%. Moreover, we find that, until the age of 45 years, a typical white male U.S. worker holds about 9 full time jobs, 50% of which are held within the first 5 years after labor market entry and roughly 75% within the first 10 years. This hints towards an enormous amount of job turnover for individuals in their first years after labor market entry. In later stages of the career job attachments become substantially more durable and job changes less frequent.

To gather insights on the observed pattern of individual job mobility we construct and estimate a life-cycle model of the labor market whose main characteristic is an information imperfection in the matching process. The key ingredient is that the imperfection is assumed to be worker-specific and in particular it is linked to an individual's previous labor market history. We estimate the structural parameters of the model by indirect inference on data from the *NLSY* 79. Using the estimates we evaluate the empirical content of our framework by assessing how well certain predictions, about individual and aggregate labor market statistics, obtained from the structural model match up with actual data. We find that the model can capture very well the observed life-cycle dimension of a variety of individual labor market outcomes. In particular it can account for the fact that job attachments tend to be very fragile early in an individual's career but become increasingly durable as workers accumulate more and more labor market experience. Furthermore, we find that the informational frictions we consider in the model are key for replicating the life-cycle profile of individual labor market outcomes.

Beyond contributing to our understanding of the observed life-cycle dynamics of job mobility, the findings of this paper have important implications, especially for the design of optimal labor market policies. Our results stress that the accumulation of work experience is key for an individual's job stability and job security. With more experience accumulated on past jobs any future employment relationship will be more stable and secure. This implies a certain path dependency which is important for labor market policies in general. Any policy that deters young workers from moving into jobs, or that prevents firms from hiring them, can be detrimental to the labor market prospects in an individual's entire career. Firing taxes or excessive employment protection are such examples that typically discourage firms from hiring young workers.

On the other hand, policies that foster the attachment of young to, and their integration into, the labor market can have significant positive employment and welfare effects as the work experience accumulated in the early years positively affects the job situation later in an individual's career. On-the-job training programs, such as the German apprenticeship program, is a good example for such a policy. The design of optimal labor market policies has to take fundamental account of this life-cycle dimension. The aggregate welfare effects of a policy must be evaluated considering not only that, at each point in time, the policy can have a different impact on young individuals than on experienced workers but also that each individual is affected differently by the policy depending on the stage of the life cycle.

References

[1] Booth, A.L., Francesconi, M. and Garcia-Serrano, 1997, Job tenure: Does history matter?, CEPR Discussion Paper No. 1531.

[2] Davis, S., and Haltiwanger, J., 1992, Gross Job Creation, Gross Job Destruction and Employment Reallocation, Quarterly Journal of Economics, 107, pp. 819-863.

[3] Farber, H.S., Hall, R. and Pencavel, J., 1993, The Incidence and Costs of Job Loss: 1982-91, Brookings Papers on Economic Activity. Microeconomics, Vol. 1, pp. 73-132

[4] Gallant, R. A. and Tauchen, G., 1996, Which Moments to Match?, Econometric Theory, 12, pp. 657-681.

[5] Gourieroux, C., A. Monfort and E. Renault, 1993, Indirect Inference, Journal of Applied Econometrics, 8, pp. S85-S118

[6] Gourieroux, C. and Monfort, A., 1996, Simulation Based Econometric Methods (Oxford: Oxford University Press).

[7] Hall, R.E., 1982, The Importance of Lifetime Jobs in the U.S. Economy, American Economic Review, 72 (4), pp. 716-24.

[8] Nagypal, Eva, 2007, Learning by Doing vs. Learning About Match Quality: Can We Tell Them Apart?, Review of Economic Studies, 74, pp. 537-566

[9] Pries, M. and Rogerson, R., 2005, Hiring Policies, Labor Market Institutions, and Labor Market Flows, Journal of Political Economy, Vol. 113(4), pp. 811-839.

[10] Smith, A. A., 1993, Estimating Nonlinear Time Series Models Using Simulated Vector Autoregressions, Journal of Applied Econometrics, 8, pp. 63 - 84

[11] Tauchen, G., 1997, New Chi-Squared Methods in Empirical Finance, in D. M. Kreps and K. F. Wallis (eds.) Advances in Economics and Econometrics: Theory and Applications: Seventh World Congress (Cambridge: Cambridge University Press) 279-317.

[12] Topel, R.H. and Ward, M.P., 1992, Job Mobility and the Careers of Young Men, Quarterly Journal of Economics, 107(2), pp. 439-79.

Appendix A

Proof of Proposition 1

First we show that the surplus function $S(\cdot)$ is linearly increasing in $\hat{\pi}$. We make use of the definition of $S(\cdot)$ as stated in 22. This gives

$$\begin{aligned}\frac{\partial S(k,i,\hat{\pi})}{\partial \hat{\pi}} = & \; y^g - y^b + \beta\rho_k(1-\sigma)\textstyle\sum_{i'\in I}\mu^e(i'|i)\left\{(1-\varphi)\frac{\partial S(k',i',\hat{\pi})}{\partial \hat{\pi}} + \varphi S^g(k',i')\right\} \\ & +\beta\rho_k(1-\sigma)\varphi\left[\textstyle\sum_{i'\in I}\mu^e(i'|i)\Gamma^u(k',i') - \Gamma^u(k',i)\right]\end{aligned} \tag{17}$$

$y^g - y^b$ is positive by definition and $S^g(\cdot) > 0$ holds whenever $y^g > b$. Furthermore, $\Gamma^u \geq 0$ for all (i,k) whenever $b \geq 0$, therefore the expression in the lower line is non-negative since $\sum_{i'\in I}\mu^e(i'|i)\Gamma^u(k',i')$ is a convex combination of $\Gamma^u(k',1)$ and $\Gamma^u(k',2)$ and we have that $\Gamma_u(k',2) \geq \Gamma^u(k',1)$. Given that it remains to find an expression for $\frac{\partial S(k',i',\hat{\pi})}{\partial \hat{\pi}}$. This can be done by recursion. Starting at the terminal date K and using $\Gamma^u(K,i) = b$ we get

$$S(K,i,\hat{\pi}) = y^g\hat{\pi}) + y^b(1-\hat{\pi}) - b \tag{18}$$

from which it follows that $\frac{\partial S(K,i,\hat{\pi})}{\partial \hat{\pi}} = y^g - y^b > 0$. Inserting this into (17) proves that $\frac{\partial S(K-1,i,\hat{\pi})}{\partial \hat{\pi}} > 0$. By recursion we thus establish that $\frac{\partial S(k,i,\hat{\pi})}{\partial \hat{\pi}}$ increasing in $\hat{\pi}$ for all k. Next we prove that $S(\cdot)$ is linear in $\hat{\pi}$. Using 17 we find

$$\frac{\partial^2 S(k,i,\hat{\pi})}{\partial \hat{\pi}^2} = \beta\rho_k(1-\sigma)(1-\varphi)\sum_{i'\in I}\mu^e(i'|i)\frac{\partial^2 S(k',i',\hat{\pi})}{\partial \hat{\pi}^2} \tag{19}$$

Applying recursion again and using $\frac{\partial^2 S(K,i,\hat{\pi})}{\partial \hat{\pi}^2} = 0$ establishes the result. To prove that $S(k,i,0) < 0$ we make use of $J(k,i,0) = (1-e)S(k,i,0)$ and show that $J(k,i,0) < 0$. We proceed by stating

$$J(k,i,0) = y^b - w(k,i,0) + \beta(1-\sigma)\rho_k(1-\varphi)\sum_{i'\in\mathcal{I}}\mu^e_{i|i'}J(k',i',0) \tag{20}$$

Notice that $J(K,i,0) = ey^b - b$ which is negative for $b > y^b$. We substitute $J(K,i,0) = ey^b - b$ into $J(K-1,i,0)$ which becomes $J(K-1,i,0) = y^b - w(K-1,i,0) + \beta(1-\sigma)\rho_k(1-\varphi)(ey^b - b)$. $J(K-1,i,0) < 0$ follows immediately from $w(K-1,i,0) > b$. Using this result we can apply recursion to find that $J(k,i,0) < 0$ for all (k,i). To prove that $J(k,i,1) > 0$ we make use of $J(k,i,1) = J^g(k,i)$, hence it remains to show that $J^g > 0$ for all (k,i). $J^g(k,i)$ is given by

$$J^g(k,i) = y^g - w(k,i,1) + \beta(1-\sigma)\rho_k(1-\varphi)\sum_{i'\in\mathcal{I}} \mu^e_{i|i'} J^g(k',i') \tag{21}$$

Notice that $J^g(K,i) = ey^g - b$ which is positive whenever $ey^g > b$. We substitute $J^g(K,i)$ into $J^g(K-1,i)$ which becomes $J^g(K-1,i) = y^g - w(K-1,i,1) + \beta(1-\sigma)\rho_k(1-\varphi)(ey^g - b)$. $J^g(K-1,i) > 0$ follows immediately from $w(K-1,i,0) < y^g$. Applying recursion to this expression we can establish the result that $J^g(k,i) = J(k,i,1) > 0$ for all (k,i).

Appendix B

The joint surplus function

The definition of the joint surplus of a match with $(k,i,\hat{\pi})$ is given by $S(k,i,\hat{\pi}) = J(k,i,\hat{\pi}) + \Gamma^e(k,i,\hat{\pi}) - \sum_{i'} \mu^b_{i|i'}\Gamma^u(k,i')$. Using the value functions (5) and (6) to substitute for $J(k,i,\hat{\pi})$ and $\Gamma^e(k,i,\hat{\pi})$ and the surplus sharing rule $\epsilon\left(\Gamma^e(k,i,\hat{\pi}) - \sum_{i'} \mu^b_{i|i'}\Gamma^u(k,i')\right) = (1-\epsilon)J(k,i,\hat{\pi})$ we arrive at the following expression for S

$$\begin{aligned} S(k,i,\hat{\pi}) = \;& E(y|\hat{\pi}) \\ & +\beta\rho_k(1-\sigma)\textstyle\sum_{i'\in I}\mu^e(i'|i)\left\{(1-\varphi)S(k',i',\hat{\pi}) + \varphi\hat{\pi}S^g(k',i')\right\} \\ & +\beta\rho_k(1-\sigma)[1-\varphi(1-\hat{\pi})]\textstyle\sum_{i'\in I}\mu^e(i'|i)\sum_{i'\in I}\mu^b(i''|i')\Gamma^u(k',i'') \\ & +\beta\rho_k\left[\varphi(1-\sigma)(1-\hat{\pi}) + \sigma\right]\textstyle\sum_{i'\in I}\mu^b(i'|i)\Gamma^u(k',i') \\ & -\textstyle\sum_{i'\in I}\mu^b(i'|i)\Gamma^u(k,i') \end{aligned} \tag{22}$$

Wage formation

Wages are determined by bilateral Nash bargaining between the worker and the firm. In particular, the wage is chosen to maximize the Nash product, hence it has to solve

$$w = \arg\max J(k,i,\hat{\pi})^e \left(\Gamma^e(k,i,\hat{\pi}) - \sum_{i'} \mu^b_{i|i'}\Gamma^u(k,i')\right)^{1-e} \tag{23}$$

where e is the fraction of the total surplus that goes to the firm. The optimality condition associated with (23) is given by

$$e\left(\Gamma^e(k,i,\hat{\pi}) - \sum_{i'} \mu^b_{i|i'}\Gamma^u(k,i')\right) = (1-e)J(k,i\hat{\pi}) \tag{24}$$

Using (a) the value functions (5) and (6) to substitute for $J(k,i,\hat{\pi})$ and $\Gamma^e(k,i,\hat{\pi})$ and

(b) the definition of the surplus function (22) we arrive at the following expression for the wage function

$$w(k,i,\hat{\pi}) = E(y|\hat{\pi}) - e\left[S(k,i,\hat{\pi}) - \beta\rho_k(1-\sigma)\sum_{i'\in I}\mu_i^e|i'\left\{(1-\varphi)S(k',i',\hat{\pi}) + \varphi\hat{\pi}S^g(k',i')\right\}\right] \quad (25)$$

Appendix C

Actual hazard rate, by job tenure and by age at match formation						
Age / Tenure	[0-1)	[1-2)	[2-3)	[3-5)	[5-7)	[7-10)
[18-20)	1.1282	0.6932	0.4868	0.3513	0.2336	0.1841
[20-23)	1.0275	0.6328	0.4431	0.3139	0.2117	0.1748
[23-26)	0.9514	0.5793	0.3990	0.2940	0.1910	0.1512
[26-30)	0.8376	0.5032	0.3563	0.2556	0.1656	0.1314
[30-34)	0.6963	0.4186	0.2915	0.2065	0.1433	0.1095
[34-38)	0.5846	0.3593	0.2459	0.1773	0.1129	0.0936
[38-	0.5362	0.3340	0.2267	0.1627	0.1054	0.0874

Estimates of workers' actual hazard rate of separating from a job for selected age and tenure groups. Source: Own calculations based on data from the *NLSY* 79.

Table 13

Duration / Age	[18-20)	[23-26)	[30-34)	[38-
1 year	0.49 **(0.57)**	0.67 **(0.63)**	0.71 **(0.72)**	0.84 **(0.79)**
2 years	0.31 **(0.43)**	0.51 **(0.51)**	0.57 **(0.60)**	0.58 **(0.67)**
3 years	0.21 **(0.37)**	0.42 **(0.43)**	0.43 **(0.52)**	0.56 **(0.59)**
4 years	0.15 **(0.29)**	0.33 **(0.37)**	0.39 **(0.45)**	0.47 **(0.52)**
5 years	0.13 **(0.25)**	0.29 **(0.33)**	0.35 **(0.41)**	0.44 **(0.46)**

Job retention rates for workers with one year of tenure in the current job: Rows represent different retention periods, columns indicate workers' age at the time of recruitment. Numbers in bold (normal font) represent the model (data).

Table 14

Appendix D: Accuracy of the estimation results

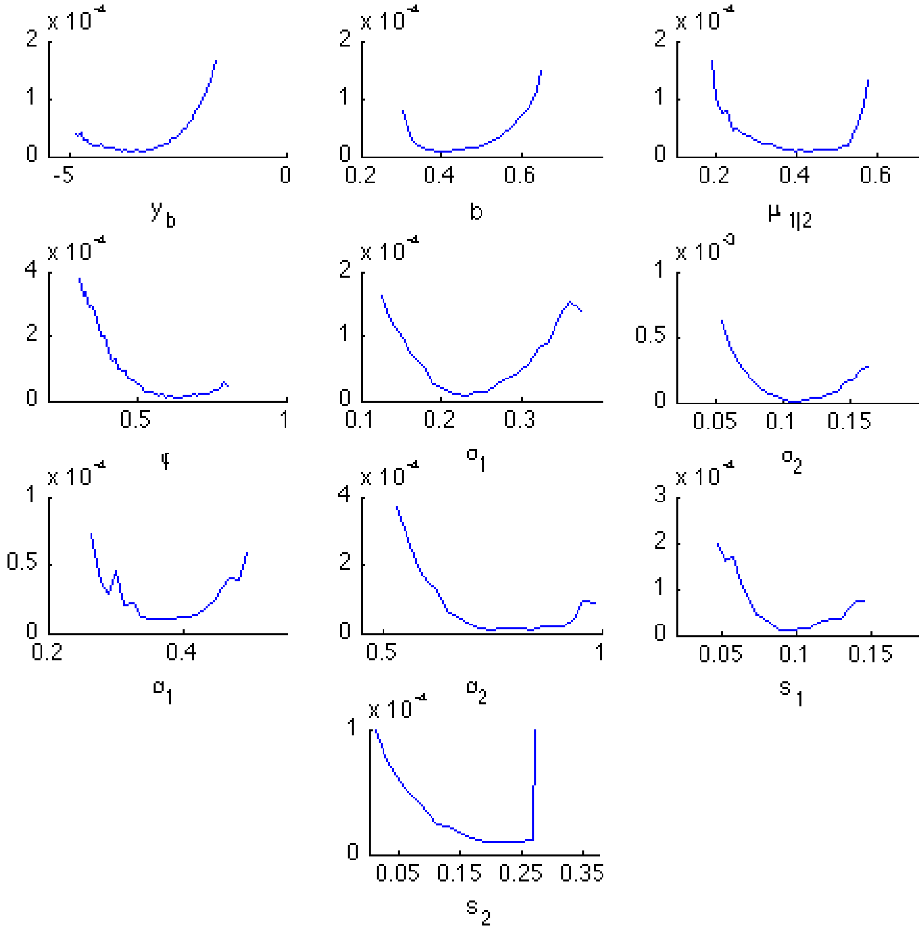


Figure 8

The value of the objective function (stated in Equation (13)) evaluated at the optimal estimator for all but one parameter. The remaining parameter is varied over a range of feasible values.