

# Returns to on-the-job search and the dispersion of wages\*

Axel Gottfries<sup>1</sup> and Coen Teulings<sup>1,2</sup>

<sup>1</sup>University of Cambridge

<sup>2</sup>University of Amsterdam

March 15, 2017

## Abstract

A wide class of models with On-the-Job Search (OJS) predicts that workers gradually select into better-paying jobs. We develop a simple methodology to test predictions implied by OJS using two sources of identification: (i) time-variation in job-finding rates and (ii) the time since the last lay-off. Conditional on the termination date of the job, job duration should be distributed uniformly. This methodology is applied to the NLSY 79. We find remarkably strong support for all implications. The standard deviation of the wage offer distribution is about 15%. OJS accounts for 30% of the experience profile, 9% of total wage dispersion and an average wage loss of 11% following a lay-off.

**Keywords:** On-the-job search, Wage dispersion, Job duration

**JEL Classification:** J31, J63, J64

---

\*Gottfries: [ag558@cam.ac.uk](mailto:ag558@cam.ac.uk). Teulings: [cnt23@cam.ac.uk](mailto:cnt23@cam.ac.uk). We would like to thank Jim Albrecht, Jake Bradley, Jan Eeckhout, Marcus Hagedorn, John Kennan, Philipp Kircher, Moritz Kuhn, Hannes Malmberg, Iouri Manovskii, Fabien Postel-Vinay, Robert Shimer and participants at NBER SI 2015, Dale Mortensen Centre Conference on Labor Market Models and Their Applications 2015, ESSLE (2015, 2016), SaM conferences in Amsterdam and Cambridge, NYU macro student lunch and Cambridge University. Gottfries thanks the Economic and Social Research Council and the Tom Hedelius Foundation for financial support. Teulings thanks the Cambridge Inet Institute, Centre for Macroeconomics and the Cooperation for European Research in Economics (COEURE) for financial support.

# 1 Introduction

Labour market models with on-the-job search (OJS) as developed by [Burdett and Mortensen \(1998\)](#) and [Bontemps, Robin, and den Berg \(2000\)](#) have become the workhorse model for explaining job-to-job transitions and wage dynamics in macro economics. By and large, these models predict an increasing and concave wage-experience profile with workers moving up the job ladder up until the point of lay-off. Some related models deviate slightly from this set up, but yield largely similar implications.<sup>1</sup>

Though this process clearly captures important features of real-life labour markets, no one has yet studied the extent to which this process adequately describes the finer details of the interrelation between job-to-job transitions and wage growth over the career of a worker. This paper fills that gap. We derive detailed predictions on both job-to-job transitions and wage growth which are then brought to the data. We find remarkably strong empirical support for this model. The returns to OJS are highly stable over the life cycle. The standard deviation of the wage offer distribution is about 15%. The return to OJS explain about 30% of the overall return to labour market experience. OJS accounts for 9% of the wage dispersion among male workers.<sup>2</sup> The standard deviation of the offer distribution is one and a half times larger for higher educated workers. For the latter group, the dispersion is one and a half times larger in urban areas rather than in the countryside. Our estimates suggest that the loss in match quality after lay-off results in an average wage loss of about 11%, compare [Jacobson, LaLonde, and Sullivan \(1993\)](#) and [Davis and Wachter \(2011\)](#).

Our methodology builds on ideas developed by [Wolpin \(1992\)](#), [Barlevy \(2008\)](#) and [Hagedorn and Manovskii \(2013\)](#). [Wolpin \(1992\)](#) introduces the concept of an *employment cycle*. An employment cycle starts at the beginning of an unemployment spell, follows the worker when he moves to his first job and subsequently to ever better jobs. The employment cycle ends when the worker gets laid off and becomes unemployed again, which starts a new cycle. In a model with OJS and efficient transitions, this corresponds to a sequence of ever-better draws from the job-offer distribution. [Barlevy](#)

---

<sup>1</sup>[Postel-Vinay and Robin \(2002\)](#) deviates on wage formation, but generates the same transition dynamics to ever "better" jobs. In [Burdett and Coles \(2003\)](#) workers do not necessarily transition to more productive jobs.

<sup>2</sup>For related decompositions, see [Manovskii and Hagedorn \(2010\)](#), [Tjaden and Wellschmied \(2014\)](#) and [Bagger, Fontaine, Postel-Vinay, and Robin \(2014\)](#).

(2008) shows how results from record theory can be used to characterise the distribution of job offers. Our approach is different. We use the information embedded in the sequence of job-durations. Our methodology has the advantage that it applies independently of the cyclical behavior of lay-off rates.

Hagedorn and Manovskii (2013) use cyclical variation in the labour market tightness to control for the effect of match quality in wage regressions. Job offers arrive more frequently when labour market conditions are tight. Hence, workers move up the job ladder more quickly. However, they do not derive the exact process. As a consequence, their regression model is misspecified. We derive the exact relation between the time elapsed since the start of the current employment cycle until the end of the current job. The starting date of a job is shown not to provide any additional information above its termination date. Moreover, it matters whether a job ends by a quit rather than a layoff: jobs that end by a quit have lower expected quality, as this raises the likelihood of the arrival of a better job.<sup>3</sup>

There are at least three reasons why wages tend to increase over the course of a worker's career: general experience, within-job tenure profiles (due to either job specific experience or a tenure specific retention bonus as in Burdett and Coles (2003)), and the gains from selection by OJS. This paper combines two sources of variation to disentangle the return to OJS from the other two sources: (i) individual variation in the timing of lay-offs: a lay-off restarts the selection of ever better offers; (ii) business cycle variation in job-offer arrival rates. While the returns to experience accumulate proportionally to calendar time, job search accumulates proportionally to what we refer to as *labour market time*. In a tight labour market, the clock of labour market time runs faster than that of calendar time, thereby speeding up the selection process. The speed of labour market time is measured by the job-offer arrival rate for unemployed.

We show that the model predicts that the starting date of the current job is uniformly distributed over the length of the employment cycle up until the termination date of this job measuring all durations in labour market time. The intuition is simple. The current job is the maximum of all job offers received over the course of the employment cycle. Since we have no further information on which of these offers is the maximum, its arrival must be

---

<sup>3</sup>See Fredriksson, Hensvik, and Nordstrom Skans (2015), Guvenen, Kuruscu, Tanaka, and Wiczer (2015) and Postel-Vinay and Lise (2015) for alternative approaches.

uniformly distributed over the employment cycle. This implication allows us to test an important assumption. If job-specific experience were important, this result would not hold, as the likelihood of a job-to-job transition would decline over the employment cycle (since, by switching jobs, a worker spoils the returns of job-specific experience accumulated so far). Surprisingly, our empirical result do not reject the null of a uniform distribution.

We show that there is a one-to-one correspondence between the expected wage of the current job as a function of the number of job offers and the shape of the offer distribution. The Gumbel distribution for log wages is not rejected against any alternative. This distribution has an unbounded upper support with a fat tail. This runs counter to models with assortative matching with an interior upper bound of the matching set (for example [Shimer and Smith \(2000\)](#) and [Gautier, Teulings, and Van Vuuren \(2010\)](#)), because in these models the offer distribution has finite upper support. Hence, sorting cannot be the only source of variation in the offer distribution.

The structure of the model of OJS in the spirit of [Burdett and Mortensen \(1998\)](#) and [Bontemps et al. \(2000\)](#) is very restrictive. This structure yields strong testable implications. We test these implications using data from the NLSY 79. We find remarkably strong support for all derived predictions. In particular i) log wages increase logarithmically by the time elapsed between the start of the current employment cycle and the moment that the current job will end. ii) this logarithmic shape is (almost) constant across all employment cycles over the life-cycle of the worker, suggesting that the offer distribution is constant over the career. iii) the two sources of variation, (a) the individual variation in the timing of lay-off shocks, and (b) cyclical variation in the job offer arrival rate, yield a similar shape of the offer distribution. iv) the model makes a specific prediction regarding the difference in expected log wages for jobs ending in a lay-off and in a quit; this prediction is supported by the data. v) the starting date of a job is irrelevant for the expected wage in that job. vi) the starting date of the current job is indeed uniformly distributed over the length of the current employment cycle. vii) the model implies that there is no within-job tenure profile in wages. Though we find some return to tenure, the estimates are small.<sup>4</sup> Point iii) and iv) imply that the empirical distinction between quits and lay-offs in the data squares well

---

<sup>4</sup>Many papers, e.g. [Dustmann and Meghir \(2005\)](#), directly estimate the tenure profile, treating the return to OJS as a residual item. This paper takes the opposite approach by providing direct estimates of the return to OJS and treating the return to tenure as the residual item

with the difference between quits versus lay-offs in the theoretical model.

The applied theoretical framework implies that workers accept any job offer that pays more than their current job. This framework fits the theoretical models of [Burdett and Mortensen \(1998\)](#), [Bontemps et al. \(2000\)](#), [Pissarides \(1994\)](#) and [Shimer \(2006\)](#) well. It also fits well the model of sequential auctions by [Postel-Vinay and Robin \(2002\)](#). However, in that model, wages are not simply an increasing function of productivity and they may change over the duration of a job. Our framework can be applied to test this characteristic, as we do in a companion paper, see [Gottfries and Teulings \(2017a\)](#). Similarly, the current paper assumes that firms can adjust wages in ongoing jobs to changes in aggregate labour market conditions. Again, our framework can be extended to test this assumption, as we do in another companion paper, see [Gottfries and Teulings \(2017b\)](#).

The structure of this paper is as follows. Section 2 develops the main theoretical concepts and derives the relation between wages and accumulated labour market experience. Section 3 presents our empirical results and section 4 concludes.

## 2 The theoretical argument

### 2.1 Assumptions

At an arbitrary time  $t$ , during their labour market career, workers receive job offers at a rate  $\lambda_t$ . A job offer is characterized by a log wage  $w$ , which is drawn from a job offer distribution function  $F(w)$ ;  $F(\cdot)$  is assumed to be differentiable;  $F(w)$  will be referred to as the rank of a job offer. Without loss of generality, we limit the offer distribution to all jobs that are acceptable to unemployed job seekers. Lower job offers are irrelevant, as they will not be accepted by any job seeker. As the labour market history of workers accumulates, they receive ever more job-offers. We assume that there are no switching cost. Hence, workers will accept any offer that pays a higher wage than their current job. Hence, the wage in their current job is the maximum of all wage offers that they have received in their employment cycle. Hence, the expected wage is increasing in the accumulated labour market history. This selection process of ever better matches continues until the worker is laid off. We assume that this happens at a, time dependent, rate  $\delta_t$ . Then, the worker becomes unemployed and the selection process of job offers starts

all over again.

This set of assumptions is consistent with the models discussed in in [Burdett and Mortensen \(1998\)](#), [Bontemps et al. \(2000\)](#), [Pissarides \(1994\)](#) and [Shimer \(2006\)](#). In [Bontemps et al. \(2000\)](#) and [Pissarides \(1994\)](#), a job offer is characterized by its productivity. However, since wages are an increasing function of this productivity in both models, we can just as well characterize a job offer by its wage rather than by its productivity. These assumption are also consistent with the model of sequential auctions by [Postel-Vinay and Robin \(2002\)](#), with one exception: workers do accept any job offer that is more productive than their current job, but wages are not an increasing function of productivity in that model. Firms can increase their wage offer at any time as to match outside offers that comes. Our methodology can be extended to allow for this type of wage setting, which we do in [Gottfries and Teulings \(2017a\)](#). Our assumptions are not consistent with the model by [Burdett and Coles \(2003\)](#), where firms pay tenure profiles as to seduce workers not to quit. In that model, workers might reject a more productive offer, since they benefit from receiving a tenure premium in their current job, while they have to accumulate tenure from scratch in their new job.

We refer to the time elapsing between two consecutive lay-offs as an employment cycle (the first employment cycle starts at the beginning of a worker's career). Hence, a worker's current employment cycle has started either at the last lay-off or -for the first employment cycle- at the start of the labour career. We normalize our measure of calendar time  $t$  such that it takes the value 0 at the start of the first job of the current employment cycle. Hence, as long as a worker has not experienced a lay-off,  $t$  is equal to labour market experience as usually defined. Note that this definition of  $t$  does not include the unemployment spell at the beginning of the employment cycle.

It is useful to define:

$$\begin{aligned}\Lambda_t &\equiv \int_0^t \lambda_r dr, \\ \Delta_t &\equiv \int_0^t \delta_r dr.\end{aligned}$$

We refer to  $\Lambda_t$  as the labour market time elapsed since the start of the first job of the current employment cycle, in contrast to  $t$ , which is the calendar time since the start of the first job. While the clock of calendar time runs at a constant rate, the clock of labour market time runs faster during a boom

(when  $\lambda_t$  is high) than during a bust (when  $\lambda_t$  is low). We define  $a$  and  $b$  as the start and end date of the current job respectively. Hence,  $\Lambda_a$  is the labour market time elapsed since the beginning of the employment cycle up until the date of the start of the current job, and  $\Lambda_b$  is the labour market time elapsed up until the date of job termination. We define  $w_a$  as the wage offer receive at the start of the current job at time  $a$ , which is therefore equal to the current wage. Finally, let  $n$  denote the number of job offers received during the current employment cycle, that is, in the time interval  $[0, b]$  (so: including the offer received at the start of the employment cycle at time 0 and at the end of the current job at time  $b$ , in case the current job ends by a quit rather than a lay-off).

We assume that the job-offer arrival rate for employed job seekers  $\lambda_t$  varies proportionally to the job-offer arrival rate for unemployed job seekers  $\lambda_{ut}$ :

$$\lambda_t = \psi \lambda_{ut}. \tag{1}$$

## 2.2 Job duration and job transition

This section derives the pattern of job duration and job-to-job transitions implied by the model. Since workers will always move to a better paying job, the only relevant statistic for the transition dynamics of workers is a job offer's rank in the offer distribution  $F(w)$ , where by construction, the rank  $F$  is normalized such that it is uniformly distributed on  $[0, 1]$ . The proposition below specifies the relation between the elapsed labour market times  $\Lambda_a$  and  $\Lambda_b$ , the number of job offers  $n$ , and the of the expected rank of the current job  $F$ .

### Proposition 1 *Transition dynamics*

1. *The expected number of job offers  $n$  in the time interval  $[0, b]$  satisfies:*  
 $E[n] = \Lambda_b + 1$  *if the job ends in a lay-off;*  
 $E[n] = \Lambda_b + 1 + O(\Lambda_b^{-1})$  *if the jobs ends in a quit.*
2. *For each job other than the first job of an employment cycle,  $\Lambda_a/\Lambda_b$  is uniformly distributed on the unit interval  $[0, 1]$ ;*
3. *For jobs ending in a lay-off, the expected rank of the current job satisfies*

$$E[F] = 1 - \Lambda_b^{-1} + O(\Lambda_b^{-2});$$

4. *For jobs ending in a quit, the expected rank satisfies*

$$E[F] = 1 - 2\Lambda_b^{-1} + O(\Lambda_b^{-2}).$$

The proof is presented in Appendix A. However, all results can be understood intuitively. The first statement says that the number of job offers until the moment of separation from the current job at time  $b$  is equal to  $\Lambda_b + 1$ . The term  $\Lambda_b$  measures the expected number of job offers after the start of the first job of the employment cycle until the moment of separation from the current job, that is, the expected number of offers in the time interval  $(0, b)$ . This number follows a Poisson distribution with arrival rate  $\Lambda_b = \int_0^b \lambda_t dt$ , that is, the arrival rate is equal to the integral over the time interval  $(0, b)$  of the arrival rate  $\lambda_t$  per unit of calendar time. Since the expectation of the Poisson distribution is equal to the arrival rate, the expected number of offers during this time interval is equal to  $\Lambda_b$ . We should add one for the job offer that yielded the first job of the employment cycle, which allowed the worker to transition from unemployment to employment at  $t = 0$ . Since the unemployed accept any job offer, we know that a job seeker transitioning from unemployment to employment has received exactly one offer at  $t = 0$ . This relation is exact for lay-offs. As will be explained in relation to statement 4, the relation applies only up until a term of order  $\Lambda_b^{-1}$  for quits, as the fact that a better offer comes at time  $b$  provides information about the number of job offers received up until time  $b$ .

The second statement states that for all jobs except for the first job of an employment cycle, the time  $a$  of the start of the current job is uniformly distributed over the labour market time from the start of the first job of the cycle at time 0 until the end of the current job at time  $b$ . The intuition for this result is that, unless a worker is still in the first job of this employment cycle, the job that a worker holds at a particular time  $b$  is the maximum of all job offers the worker has received in the time interval  $(0, b)$ . Since job offers arrive proportional to labour market time and since conditional on the total number of offers received over this time interval, there is no prior information about which of these offers is the highest offer, the arrival is uniformly distributed over labour market time in this time interval.

The third and the fourth statements are about the expectation of the rank of the current job, conditional on the labour market time elapsed until the end of that job. This expectation depends on the reason for separation from the current job: either a lay-off or a quit. First, consider the case of



separation by means of a lay-off (see statement 3). Then, the rank of the current job is the max of  $\Lambda_b + 1$  expected draws from the job-offer distribution (see statement 1). The expected maximum of  $n + 1$  draws from the uniform distribution is  $1 - (n + 2)^{-1}$ ,<sup>5</sup> which explains the result. The relation applies only up to a term of order  $\Lambda_b^{-1}$ , because we replace the actual number of job offers  $n + 1$  by its expectation,  $\Lambda_b + 1$ . By Jensen's inequality,  $E[(n + 1)^{-1}] > E[n + 1]^{-1}$ . The difference is of order  $\Lambda_b^{-2}$ . Note that the difference between  $(\Lambda_b + 2)^{-1}$  and  $\Lambda_b^{-1}$  is also of order  $\Lambda_b^{-2}$ ; hence, this difference is encompassed in the term  $O(\Lambda_b^{-2})$ .

Next, consider the case of separation from the current job by means of a quit (see statement 4). An easy way to calculate the max is to consider the number of offers received up until time  $b$ , including the offer of the job to which the worker transitions at time  $b$ . The number of offers is  $n + 2$ . In this case, the job that the worker holds up until time  $b$  is not the maximum offer, but the second highest, since the offer to which the worker transits at time  $b$  is higher (otherwise the worker would not have moved to that job). The expectation of the second to highest job offer after  $n + 2$  offers is  $1 - 2(n + 2)^{-1}$ . Again applying  $E[(n + 2)^{-1}] = E[n + 2]^{-1} + O(\Lambda_b^{-2}) = \Lambda_b^{-1} + O(\Lambda_b^{-2})$  yields the result. Note that in both statements 3 and 4,  $E[F] \rightarrow 1$  if  $\Lambda_b \rightarrow \infty$ : if the selection of ever-better offers is allowed to continue forever, the actual rank will converge to maximum rank,  $F = 1$ .

Proposition 1 provides a framework for understanding the methodology to test the model discussed in Section 2.1. A particularly attractive feature of Proposition 1 is that none of its statements depend on the lay-off rate  $\delta_t$  and its integral over the time interval  $[0, t]$ ,  $\Delta_t$  and hence that we do not have to worry about the covariation of  $\lambda_t$  and  $\delta_t$ . The intuition for this result is that a lay-off only stops the selection process in the current employment cycle. It does not interfere with the selection process itself. Statement 3 implies that the expected rank of the current job  $E[F]$  depends on its termination date  $b$ , but does not depend on its starting date  $a$ . Hence, we can ignore the value of  $a$  in our analysis of the effect of OJS on log wages. Hagedorn and Manovskii (2013) use a concave (logarithmic) transformation of labour market tightness elapsed over the interval  $[0, a]$  and over the interval  $[a, b]$  (that is: they use  $\ln \Lambda_a$  and  $\ln(\Lambda_b - \Lambda_a)$ ) to control for match quality in their wage regression. These measures are not sufficient statistics for the expected match quality,

---

<sup>5</sup>The expectation of the  $i$ -th order statistic for  $n$  draws of a uniform distribution is  $i/(n + 1)$ .

since  $\ln \Lambda_b$  is not a linear combination of  $\ln \Lambda_a$  and  $\ln (\Lambda_b - \Lambda_a)$ .

Statement 3 allows an analysis of the potential role of job specific-experience in the model. For the sake of the argument, let us assume that on-the-job experience has a linear impact on  $w_t$ :

$$w_t = w_a + \beta_x (t - a).$$

It is easy to see that the optimal strategy of the worker is no longer to quit at time  $b$  for any job for which  $w_b > w_a$ . Instead, the optimal strategy is to quit if  $w_b > w_a + \beta_x (b - a)$ . This statement can be generalized. Let  $S$  be the set of arrival times  $s$  of new job offers during the current employment cycle. Then, the job offer  $w_a$  currently held by the worker satisfies

$$w_a = \arg \max_{s \in S} [w_s + \beta_x (a - s)].$$

The worker is prepared to move to a better job only if the gain in  $w_a$  offsets the loss in job-specific experience in the previous job. The selection process can still be described as the max over a number of draws from an offer distribution, but the offer distribution is non-stationary: it gradually deteriorates at a rate  $\beta_x$  per unit of calendar time. Hence, labour market time spent early in the career is more valuable for the selection process than time spent later as job offers at that stage are more attractive since they leave a longer time period to accumulate job-specific experience. This implies that workers will change jobs more often early in their career than is predicted by statement 2. If the distribution of  $\Lambda_a/\Lambda_b$  is skewed to the left, this shows that job-specific experience plays a role. Hence, the distribution of  $\Lambda_a/\Lambda_b$  provides a test for the relevance of job-specific experience. The equations are more complicated for the case with, rather than without, job-specific experience. For the sake of transparency, we therefore focus the subsequent discussion on the case without job-specific experience. Extending the theory to the case with job-specific experience is straightforward, in principle.

## 2.3 Measuring match quality

### 2.3.1 Identifying the offer distribution

Thus far, we have made no assumptions regarding the shape of the offer distribution except for the differentiability of  $F(w)$ . This section discusses how

information of the conditional expectations can be used to identify the distribution. Proposition 2 states that the function  $E[w_a|\Lambda_b, \text{Lay off}]$  identifies the distribution  $F(w)$ .

**Proposition 2 *Non-parametric identification*** *If the expected maximum from  $n$  iid draws from an distribution  $F(w)$  is finite for all  $n$  then the function  $E[w_a|\Lambda_b, \text{Lay-off}]$  non-parametrically identifies the distribution  $F(w)$ .*

The proof is presented in Appendix B.  $F(w)$  can be non-parametrically identified from entry wages, or the steady state distribution of wages, if wages are perfectly measured (see Bontemps et al. (2000)). Barlevy (2008) shows that the sequence of wage changes at job moves (sequence of record changes: from the first to the second job, from the second to the third etc) characterises the distribution up to the location parameter. In practice only the first few wage changes can be precisely estimated as there are few employment cycles with more than 4 jobs. Our approach uses another source of information, namely variation in the length of the job duration measured in labour market time (the record time in Barlevy (2008)) to estimate the distribution more precisely. Though in theory the offer distribution is fully identified and though our method uses more information than Barlevy (2008), this information is still insufficient for a full identification of the distribution in practice. In the next section we show how we can nevertheless make strong statements about the class of distributions that is consistent with the observed pattern of job-to-job transitions and wage growth observed in the data.

### 2.3.2 Measuring the expected match quality for high values of $b$

For sufficiently large values of  $b$  we can make strong predictions regarding the shape of  $E[w_a|\Lambda_b, \text{Lay off}]$  without knowing the exact distribution  $F(w)$ . A high value of  $b$  implies that the expected number of job offers received since the beginning of the employment cycle is large. In that case we can invoke Extreme Value Theory. The normalised maximum of a large number of draws from a from a broad class of distributions converges to the Generalized Extreme Value (GEV) distribution. Denote the maximum  $w_n$  of  $n$  i.i.d. draws from some distribution  $F(w)$  and let  $\mu_n$  and  $\sigma_n$  be some normalizing constants. If the distribution of  $\frac{w_n - \mu_n}{\sigma_n}$  converges to a stable distribution, then this distribution is the Generalized Extreme Value distribution. Intuitively, the constants  $\mu_n$  and  $\sigma_n$  ensure that the mean and variance respectively are

well-behaved. Theorems 1.13 and 1.2.1 in [De Haan and Ferreira \(2007\)](#) prove the relevant results.

The GEV distribution has three parameters: (i)  $\xi$  the shape parameter, (ii) the scale parameter  $\sigma$  and (iii) the location parameter  $\mu$ . The cases  $\xi < 0$ ,  $\xi = 0$  and  $\xi > 0$  correspond to the Weibull, Gumbel and Frechet distribution, respectively. If the normalized maximum of a distribution converges to the GEV distribution with  $\xi < 1$  then the expectation of the normalised maximum converges to the expected value of the GEV distribution (see Theorem 5.3.1 in [De Haan and Ferreira \(2007\)](#)). Using the normalising constant  $\sigma_n$  and  $\mu_n$  we can find a measure of the expected match quality for large values of  $n$ . In particular since  $\frac{E[w_n] - \mu_n}{\sigma_n} \approx E[x]$ , where  $E[x]$  denotes the expectation of the extreme value distribution, the following expression is a measure of the expected match quality

$$E[w_n] \approx \mu_n - \sigma_n E[x] \quad (2)$$

The speed of convergence to the GEV distribution differs across distributions; it is  $n^{-1}$  for the exponential distribution, whereas it is much lower for the normal distribution, only at a rate  $(\ln n)^{-1}$ . The speed of convergence depends on the shape of the right tail of the distribution. For example, after just four job offers,  $\Pr(F < 0.5) = 0.0625$  and  $E[F] = 0.80$ . The left tail of the distribution becomes irrelevant even for a low number of job offers and only the upper deciles of the distribution matter. If these deciles fit the Weibull, Gumbel or Frechet distribution well, then convergence will be fast.

If  $F$  is the GEV distribution then the relation is exact and the expected value takes a particularly simple form

$$E[w_n] = \begin{cases} \mu + \gamma + \sigma \ln n & \text{if } \xi = 0 \\ \mu + \sigma \xi^{-1} [\Gamma(1 - \xi)] n^\xi - 1 & \text{if } \xi < 1 \vee \xi \neq 0, \\ \infty & \text{if } \xi \geq 1 \end{cases} \quad (3)$$

$$\text{Var}[w_n] = \begin{cases} \frac{\pi^2}{6} \sigma^2 & \text{if } \xi = 0 \\ \frac{\Gamma(1-2\xi) - \Gamma(1-\xi)^2}{\xi^2} \sigma n^{2\xi} & \text{if } \xi < \frac{1}{2} \vee \xi \neq 0 \\ \infty & \text{if } \xi \geq \frac{1}{2} \end{cases} \quad (4)$$

where  $\gamma = 0.577$  is Euler's constant. The variance decreases with  $n$  if  $\xi < 0$ , increases if  $\xi > 0$ , and remains constant if  $\xi = 0$ . Similarly, the

expectation of  $w_n$  per unit increase in  $\ln n$  decreases with  $n$  if  $\xi < 0$ , increases if  $\xi > 0$ , and remains constant if  $\xi = 0$ . Intuitively, it is relatively quickly to climb up a thin tailed distribution whereas it takes a long time if the distribution is fat tailed. When  $F(w)$  follows the Gumbel distribution,  $F(w) = \exp(-e^{-(w-\mu)/\sigma})$ , the transformation of  $F(w)$  to the GEV distribution  $G(x)$  is particularly simple since  $\mu_n = \mu + \sigma \ln n$  and  $\sigma_n = \sigma$ .

The formulas presented are for the case when the number of offers is fixed. In our case, this number is stochastic, following a Poisson distribution. We use the notation  $\cong$  whenever the approximation  $E[f(n)] \cong f(E[n])$  is used<sup>6</sup>. Propositions 1 and equation 3 can be used for the derivation of expressions for  $E[w_a|b, \text{lay-off}]$  and  $E[w_a|b, \text{quit}]$ . When a job ends with a lay-off,  $w_a$  is the maximum of  $n + 1$  offers and  $E[n|\Lambda_b] = \Lambda_b$ ; see Proposition 1. The expected value of  $w_a$  is the max of an expected number of offers equal to  $\Lambda_b$  satisfies

$$E[w_a|b, \text{lay-off}] \cong \mu + \sigma \ln(\Lambda_b + 1) \quad (5)$$

Figure 1 plots the expected maximum for the case of a fixed number of offers of  $\Lambda_b + 1$  compared to the case where the number of offers follows a Poisson distribution. We plot the GEV distribution for values of  $\xi \in \{-0.4, 0, 0.4\}$  and the Normal distribution. The location and scale parameter for each distribution is chosen such that the underlying job-offer distribution  $F(w)$  has a zero expectation and a unit variance. The figure shows that the difference between a fixed number of offers and a Poisson distributed number is small relative to the effect of the shape of the offer distribution.

When a job ends by a quit, the distribution of  $w_a$  is the the second highest draw from  $n$  draws. The second highest draw is in turn given by the linear combination  $nE[w_a|n-1] - (n-1)E[w_a|n]$  of the highest draw. Using the approximation for the number of offers from Proposition 1, statement 4, and the approximation  $n \cong E[n]$  and the Gumbel distribution yields an appealing result<sup>7</sup>:

$$E[w_a|b, \text{quit}] \cong \mu + \sigma \ln(\Lambda_b + 1) - \sigma + O(\Lambda_b^{-1}). \quad (6)$$

---

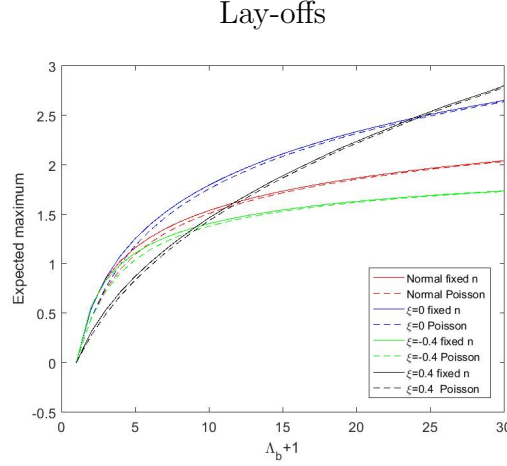
<sup>6</sup>The exact formula applies the expectation

$$E[w_a|b] = \sum_0^\infty \Pr(n|b) E[w_a|n] = \mu + \sum_0^\infty (n!)^{-1} \Lambda_b^n e^{-\Lambda_b} \sigma \ln(n+1).$$

The formula in the text uses the first-order expansion  $E[w_a|b] \cong E[w_a|E[n]]$ .

<sup>7</sup>Where we use that  $n \ln\left(\frac{n}{n-1}\right) \sigma = -\sigma + O(n^{-1})$ .

Figure 1: Expectation of the maximum as a function of the number of draws for four distributions



For large values of  $\Lambda_b$  the difference between quits and lay-offs is only a location shifter, where the size of the shift is equal to  $\sigma$ . Hence, we combine data on  $\Lambda_b$  for jobs ending in a quit or lay-off by including a dummy for jobs that end in quits. The coefficient for this dummy should be equal to  $\sigma$ .

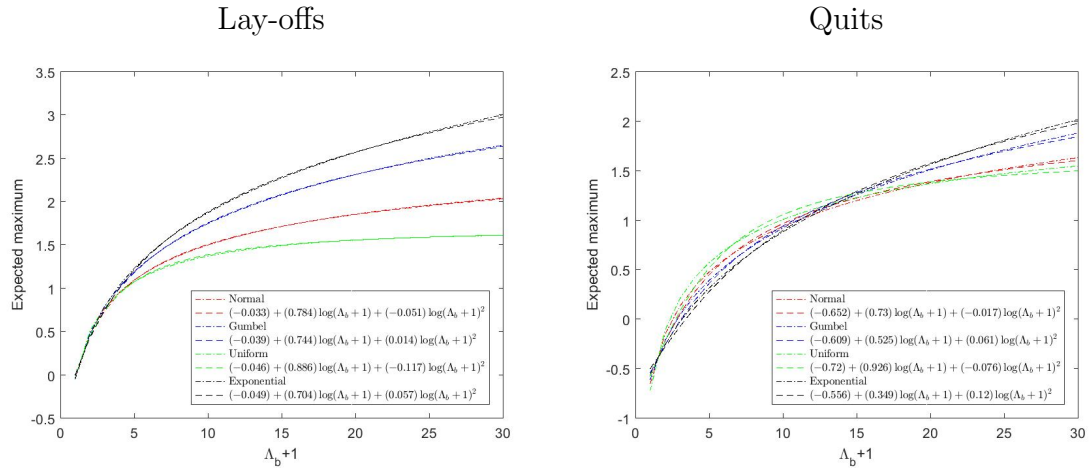
### 2.3.3 Measuring the expected match quality for any value of $b$

As we discussed before, distributions such as the Normal distribution converge slowly to the GEV distribution. For such distributions using the approximations from extreme value theory provides little guidance for the values of  $\Lambda_b$  that are relevant in practice. An alternative approach is to fit a flexible function of  $\Lambda_b$  to the conditional expectation. We illustrate this approach for a range of distributions. For values of  $\Lambda_b + 1 \leq 30$ , the expectation is approximated well by a second-order polynomial in  $\ln(\Lambda_b + 1)$ . Figure 2 plots the expected value for lay-offs and quits as a function of  $\Lambda_b$  for different distributions and the predicted values based on the regression using  $\beta_0 + \beta_1 \ln(\Lambda_b + 1) + \beta_2 \ln(\Lambda_b + 1)^2$ .<sup>8</sup> The approximation of the expected match quality fits the conditional expectation for the different distributions

<sup>8</sup>We use  $3 \times 10^6$  random draws for each value of  $\Lambda_b$  to calculate the conditional expectations. The conditional expectation is then regressed on  $\beta_0 + \beta_1 \ln(\Lambda_b + 1) + \beta_2 \ln(\Lambda_b + 1)^2$

very well. This is important, as it suggests a simple test for the null hypothesis whereby the data is generated from a particular distribution. Each distribution implies values for  $\beta_1$  and  $\beta_2$ , which can be derived numerically. We can then test the actual coefficients against any null hypothesis.

Figure 2: Approximation of the expected maximum



### 3 Empirical analysis

#### 3.1 Data

We use the cross-sectional sample from NLSY79 over the years from 1979 to 2012. Since many women interrupt their working career for childbearing, a phenomenon that is not covered in our theoretical model, we focus on males. Similarly, since our model applies to primary jobs, the sample is restricted to the primary jobs for men over the age of 18 who are not enrolled in full-time education.<sup>9</sup> We exclude job observations in cases when hours worked per week are less than 15 and when job spells lasted shorter than four weeks

<sup>9</sup>Enrollment is not recorded in some waves in 2008, 2010 and 2012 but at this point in the sample the respondents were in their 40s and few were enrolled in the previous waves.

or started before 1979.<sup>10</sup> When there are multiple jobs, the primary job is defined as the job with highest number of hours.<sup>11</sup> Jobs with inconsistencies in their start and end date are adjusted or removed.<sup>12</sup> If schooling is not reported for a given month, we assign the maximum from the previous months; if it is less than previously reported, we use the max previously reported.<sup>13</sup>

For the construction of the variable  $\Lambda_b$ , we have to categorise job terminations into either quits (belonging to the same employment cycle) or lay-offs (starting a new cycle). We follow Barlevy (2008), who define a separation as a quit when the new job starts within eight weeks of the termination of the previous job and the stated reason for separation was voluntary (where a non-response is treated as voluntary).<sup>14</sup> If two jobs overlap, we consider the transition to be voluntary if the last job is the primary job, over the overlapping period. Jobs that begin as non-primary jobs and then become primary jobs are dropped, as are all jobs following in the employment cycle. This definition is used to determine whether or not two consecutive jobs belong to the same employment cycle.

Having defined employment cycles, we have to decide which jobs to include in our analysis of jobs. We exclude jobs which have not ended. Jobs end if the worker reports that he no longer works at the job, if the job becomes a secondary job or if the worker at an interview during the subsequent year

---

<sup>10</sup>If information on the number of hour's is not available for an observation in a job spell we assigned the average over that job spell.

<sup>11</sup>If hours are the same, the average hours for that job spell, the length of the job spell and then the wage are used to determine which is the main job. 40 observations are dropped where the value, for all these variables, is the same as another observation for the individual.

<sup>12</sup>Observations missing information on the month or year when the job started or ended were removed from the data. If the day is unknown we set it to 15. If the day reported is greater than the number of days in the month (e.g. 31st of February), we set the day to the last day of the month. If at the interview the worker reported that the job ended after the interview date, we set the end date to the interview date. Jobs where the start date is reported as being after either the interview date or the end date are dropped.

<sup>13</sup>We use the "adjusted" schooling variable.

<sup>14</sup>We deviate in our approach for cases where the worker stated "leaving to look for another job" . We consider a job change a quit only if the next job starts within two weeks. A worker might have had an outside offer but quit his job to look for an even better offer. Alternatively, the worker's requirements for the job might have changed (e.g. the worker has to move to another city). The first example should be classified as a quit (hence as a continuation of the current employment cycle), while the second ends the employment cycle.



does not mention working for the firm during the past year. Jobs where the worker reports being self-employed or working for a family business, or where the hourly wage is below \$1 or above \$500, or where some of the covariates are missing values are dropped from the analysis. Wages are deflated using seasonally adjusted national CPI (CPIAUCSL).

We calculate the transition rates using the monthly CPS data. We restrict our analysis to a sample of males age 25-54 in order to match our NLSY dataset and avoid moves involving voluntary participation decisions as opposed to job-offer arrivals.<sup>15</sup> To calculate the job-finding rate of the unemployed, we calculate the fraction of the workers who are unemployed less than five weeks and are employed in the next month.<sup>16</sup> In appendix C we discuss how we can estimate of  $\psi$  from the observed value of  $b$  in the first and subsequent jobs. The variable  $\Lambda_b$  is the created as the sum of the monthly job finding rate scaled by  $\psi$ . We use the non-seasonally-adjusted unemployment rate for men 25-54 created by BLS (LNU04000061). Table 1 provides summary statistics for the variables of interest.

## 3.2 The distribution of job tenure

We derive an estimate for  $\psi$  from the difference in the value of  $\Lambda_b$  for the first job in an employment cycle and for all subsequent jobs, the method is discussed in Appendix C. From this we can derive estimates for the average value of  $\delta_t$  and  $\lambda_t$  over all  $t$ ;  $\psi$  is equal to the ration of this calculated average value of  $\lambda_t$  and the observed average value of  $\lambda_{ut}$ . The latter is observed from the monthly transitions rates from unemployment to employment in the CPS monthly data. In table 2 we show the value of  $\Lambda_b$  in the first and subsequent jobs together with the estimated average transition rates.

Statement 2 of Proposition 1 implies that for all jobs except the first job of an employment cycle the labour market time at the moment of the start of a job should be uniformly distributed over the employment cycle up until the termination date of the job. This implies that  $\Lambda_a/\Lambda_b$  for all jobs other than the first job of an employment cycle should be distributed

---

<sup>15</sup>We match the monthly CPS using variables suggested by [Drew and Warren \(2014\)](#). In addition, we use race and age as extra controls.

<sup>16</sup>Due to changes to the CPS classification, the monthly files cannot be matched for a small number of months (07/1985, 10/1985, 01/1994, 06/1995, 07/1995, 08/1995 and 09/1995). For these months we use the predicted values from a regression of the transition rate on a linear trend, a monthly fixed effect and the current unemployment rate.

Table 1: Summary Statistics

	First Job	Second Job	Subsequent Jobs	Total
Fraction urban	0.801 (0.416)	0.795 (0.415)	0.795 (0.426)	0.789 (0.426)
Fraction high Edu.	0.316 (0.465)	0.394 (0.489)	0.444 (0.497)	0.359 (0.480)
ln real hourly wage	-2.841 (0.540)	-2.656 (0.546)	-2.532 (0.583)	-2.681 (0.558)
$\ln(\Lambda_b + 1)$	0.709 (0.598)	1.300 (0.690)	1.797 (0.668)	1.345 (0.879)
$\ln(\Lambda_{Tb} + 1)$	1.569 (0.837)	1.952 (0.666)	2.190 (0.582)	1.968 (0.803)
$\ln(\lambda b + 1)$	0.716 (0.596)	1.307 (0.682)	1.803 (0.660)	1.352 (0.874)
Individuals	2572	1470	607	2582
Jobs	12623	2254	991	15868
Observations				33386

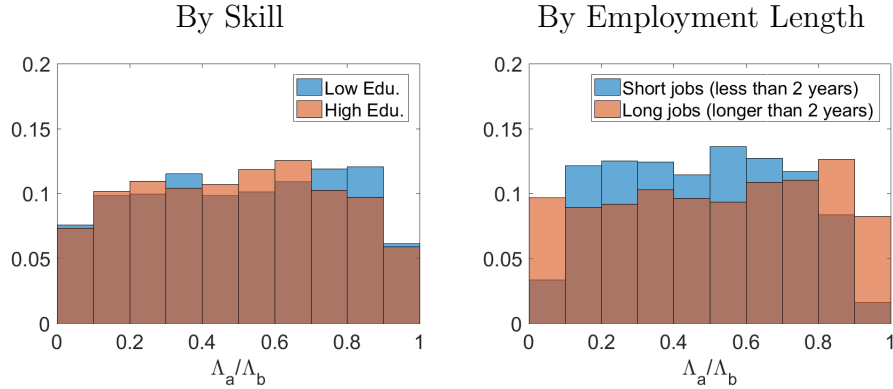
Notes: For the columns First job, Second job and Subsequent jobs, only the first observation for each job is used. For the Total column, all observations are included. Standard deviations are in parentheses. High Edu. refers to more than 12 years of education.  $T_b$  refers to the sum of all previous employment cycles up until the end of the current job. Individuals refers to the number, in our sample, without missing values of individuals.

Table 2: Transition rates

$E[\Lambda_b   1^{st} \text{ job in emp. cycle}]$	=	1.6,
$E[\Lambda_b   \text{subsequent jobs}]$	=	4.6.
$\lambda_u$	=	0.40
$\lambda$	=	0.08
$\delta$	=	0.02

uniformly. Figure 3 presents the histogram of  $\Lambda_a/\Lambda_b$ , separately for low- and highly educated workers. The actual distributions fit the uniform distribution remarkably well. If job-specific human capital had played a major role, the distribution would have been downward sloping. There is no evidence of that, except for the last decile of the distribution, where the density function falls sharply. The latter can be explained by the fact that we ignore jobs lasting less than four weeks. For example, if the employment cycle at the termination date of the current job has lasted for two years,  $\Lambda_a/\Lambda_b$  can never be above  $23/24 = 0.958$ . To test this more formally, we run the test separately for  $b < 2$  years and  $b \geq 2$  years; see Figure 3. The sample for  $b \geq 2$  year exhibits a much smaller decline of the density function, and the decline starts at a higher point in the distribution. Performing a Kolmogorov-Smirnov test for the null hypothesis on all data we reject the null of a uniform distribution the 1% level. If we restrict the sample to  $\Lambda_a/\Lambda_b \in [0.05, 0.95]$  in order to deal with the censoring at the extremes of the distribution, we can not reject null hypothesis at the 10% level.

Figure 3: Test of the arrival rate of the maximum



These results provide a strong and unexpected confirmation of the simple theoretical model of OJS that we apply. We had expected that job-specific human capital would play an important role, leading to a negative duration dependence in the job-to-job transition rate, even after controlling for the initial match quality  $w_a$ . Our results show the opposite.

### 3.3 The shape of the offer distribution

The uniform distribution of  $\Lambda_a/\Lambda_b$  justifies the application of Proposition 3 for the analysis of the evolution of the match-specific component in log wages over the employment cycle. This section presents the tests of the model as well as an estimate for the shape of the offer distribution. Worker  $i$ 's log wage at time  $t$ , denoted  $\bar{w}_{it}$ , satisfies

$$\bar{w}_{it} = \beta_i + \beta' X_{it} + w_{ia} + \varepsilon_{it}, \quad (7)$$

where  $\beta_i$  is a worker fixed effect measuring unobserved general human capital,  $X_{it}$  is a vector measuring observed general human capital obtained by either education or work experience,  $a$  is the time at which the worker started in her current job,  $w_{ia}$  measures the component of wages that is specific to the current job, and  $\varepsilon_{it}$  is a random variable. The three components  $\beta_i + \beta' X_{it}$ ,  $w_{ia}$  and  $\varepsilon_{it}$  are mutually uncorrelated. For the moment, we abstract from job-specific work experience. Taking expectations conditional on  $\Lambda_b$  we obtain the expression

$$\bar{w}_{it} = \beta_i + \beta' X_{it} + E[w_{ia}|\Lambda_b] + \nu_{ia} + \varepsilon_{it}, \quad (8)$$

where  $\nu_{ia} \equiv w_{ia} - E[w_{ia}|\Lambda_b]$ . We estimate the above equation where we include in the vector  $X_{it}$  tenure and experience up to a third order (with interactions) and quadratics in years of education and time (measured in months), as well as dummies for region, marriage, and urban versus rural location. We add the unemployment rate as a proxy for the effect of general labour market conditions on wages as well as a dummy for jobs that end in a quit in order to correct for the difference in expected job quality for jobs ending by a quit or a lay-off (see Proposition 1). Standard errors are clustered at the job level.

Our model is inconsistent with a return to tenure, since it would imply that the match-specific term  $w_{ia}$  is non-constant over the duration of the job. If there would be a deterministic tenure profile in wages as in [Burdett and Coles \(2003\)](#), then the optimal strategy of workers would no longer be to accept each job offering a higher value of  $w_{ia}$  than in the current job. The results on the distribution of  $\Lambda_a/\Lambda_b$  presented in the previous section suggest that job-specific returns are irrelevant. Nevertheless, we include this third-order polynomial to test this prediction. As discussed in the introduction of the paper, our framework can be extended to test for the presence of a stochastic return to tenure that which characterizes the sequential auction

Table 3: Test for the measure of match quality

	All Data	Low-Edu.	High-Edu.
$u_t$	-0.011*** (0.002)	-0.014*** (0.002)	-0.008*** (0.003)
$\ln(\Lambda_b + 1)$	0.116*** (0.012)	0.106*** (0.015)	0.123*** (0.021)
$\ln(\Lambda_b + 1)1(First)$	-0.023* (0.012)	-0.024 (0.015)	-0.005 (0.021)
$\ln(\Lambda_a)$	0.002 (0.007)	-0.012 (0.008)	0.020* (0.011)
First job	0.016 (0.019)	0.023 (0.023)	-0.015 (0.034)
Observations	33386	20885	11722
$R^2$	0.645	0.565	0.675

Standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ 

model by [Postel-Vinay and Robin \(2002\)](#). We do this in a companion paper, see [Gottfries and Teulings \(2017a\)](#).

Conditional on  $\ln(\Lambda_b + 1)$  and if it is the first job the start date of the current job provides no further information. In table 3 we test this prediction by including  $\ln(\Lambda_a)$ , with value zero for first jobs. The coefficient on  $\ln(\Lambda_a)$  is small and only significant at the 10% level.

Proposition 2 shows that the function  $E[\Lambda_b, \text{Lay-off}]$  non-parametrically identifies the distribution. We create 20 bins based on the value of  $\ln(\Lambda_b + 1)$  such that each has equal number of observations. We create dummy variables for each bin and regress the log wage on the controls and the dummy variable. The coefficients, together with estimated standard deviations, for the different dummies from the regressions are presented in Figure 4. The results show that the conditional expectation is approximated well by a linear function in  $\ln(\Lambda_b + 1)$  as would be the case if the data were generated by the Gumbel distribution.

We perform a separate test for the Normal, Gumbel, Uniform and Exponential distribution. From Figure 2 we take the coefficients from the poly-

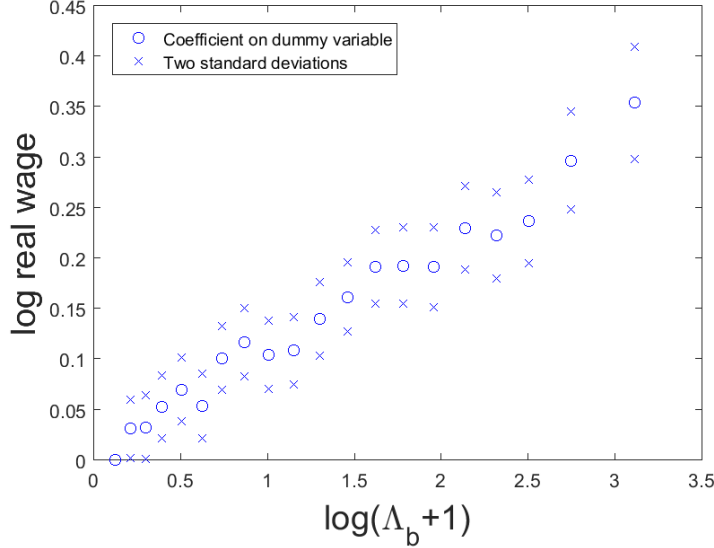


Figure 4: Test for the shape of the distribution

nomial in  $\ln(\Lambda_b + 1)$  to create the conditional expectation of the distribution denoted  $E[w_a | \text{Dist}, \Lambda_b]$ . We then run a regression using  $E[w_a | \text{Dist}, \Lambda_b]$  and  $\ln(\Lambda_b + 1)^2$  as explanatory variables. Under the null hypothesis, that the data came from a particular distribution, the coefficient of  $\ln(\Lambda_b + 1)^2$  should be zero. The results appear in table 4. The null-hypothesis of the offer distribution being uniformly distribution is clearly rejected whereas the other distributions are not rejected.

The rejection of the uniform distribution has an important theoretical implication. It rules out pure sorting models with assortative matching, where matching sets are convex in the type space and the best match is an interior maximum over this set (e.g. [Shimer and Smith \(2000\)](#) and [Gautier et al. \(2010\)](#)). An interior maximum implies that the upper support of the distribution should be finite. The estimation results are inconsistent with this prediction. Though there is empirical support for this type of sorting (see [Gautier and Teulings \(2015\)](#)), sorting cannot be the full story. In all subsequent regressions, we use the Gumbel distribution and hence logarithmic transformation of  $\Lambda_b + 1$ .

Table 4: Test for the distribution

	Normal	Gumbel	Uniform	Exponential
$u_t$	-0.012*** (0.002)	-0.012*** (0.002)	-0.012*** (0.002)	-0.012*** (0.002)
$E[w \text{Dist}, \Lambda_b]$	0.141*** (0.019)	0.138*** (0.021)	0.117*** (0.016)	0.117*** (0.022)
$\ln(\Lambda_b + 1)^2$	0.005 (0.004)	-0.001 (0.006)	0.012*** (0.003)	0.001 (0.006)
Observations	33386	33386	33386	33386
$R^2$	0.645	0.644	0.645	0.644

Standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Notes: Regressions of the logarithm of the real wage on the conditional expectation for different distributions.

### 3.4 Further tests of the OJS model

This section presents some tests of the detailed implications of the model for the evolution of wages over the career of a worker, maintaining the assumption that job offers are generated by a Gumbel distribution. All regressions that we run for this purpose are done separately for low- and highly educated workers. Across all regressions, the effect of OJS tends to be one and half times bigger for high- rather than low-educated workers. For the sake of comparison with the literature, we report the value of the coefficient for the unemployment  $u_t$  for all regressions. This coefficient is remarkably stable across all specifications and all subgroups at a value of about  $-0.01$ : one percentage point increase in unemployment reduces real wages by one percent.

Our first test checks whether the coefficient  $\sigma$  for  $\ln(\Lambda_b + 1)$  is stable across subsequent employment cycles, as is predicted by the model. The table provides a test for the role of the two different sources of identification. In Table 13, this coefficient is allowed to be different for the different employment cycles, e.g. the first cycle of the career starting at labour market entry versus subsequent cycles after lay-offs. The results provide strong support for the prediction that  $\sigma$  is stable across employment cycles. Though the coefficient for  $\ln(\Lambda_b + 1)$  is somewhat lower for later cycles, the order

Table 5: Estimation for the different employment cycles

	All Data	Low-Edu.	High-Edu.
$u_t$	-0.012*** (0.002)	-0.014*** (0.002)	-0.008*** (0.003)
$\ln(\Lambda_b + 1)1(\text{cycle} = 1)$	0.119*** (0.011)	0.093*** (0.013)	0.150*** (0.018)
$\ln(\Lambda_b + 1)1(1 < \text{cycle} \leq 4)$	0.118*** (0.006)	0.088*** (0.007)	0.149*** (0.010)
$\ln(\Lambda_b + 1)1(4 < \text{cycle} \leq 7)$	0.103*** (0.007)	0.095*** (0.009)	0.115*** (0.014)
$\ln(\Lambda_b + 1)1(7 < \text{cycle})$	0.066*** (0.009)	0.071*** (0.010)	0.072*** (0.019)
Quit	-0.042*** (0.007)	-0.056*** (0.008)	-0.014 (0.013)
Observations	33386	20885	11722
$R^2$	0.645	0.565	0.676

Standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

of magnitude is the same and the coefficient is highly significant for each cycle. This is a strong confirmation of our model. In Appendix D we present separate estimates for each of the first eight cycles. These results yield the same conclusion.

Table 13 provides a further test of the model. Jobs ending by a quit should pay less than jobs ending by a lay-off, since the arrival of a better offer (which is required for a worker willing to quit) signals that —other things equal— the current job has a relatively low wage. Our theoretical derivation shows that this affect is of order  $\sigma$ . Our results show indeed that jobs ending in a quit pay less, though the effect smaller is than  $\sigma$ .

For our next test we check whether the coefficient  $\sigma$  for  $\ln(\Lambda_b + 1)$  is stable across experience, as is predicted by our model. We interact the measure of match quality with the level of experience to test if the scale parameter is stable across the career. The test does not reject the null when we combine



Table 6: Estimation by level of experience and age

	All Data	Low-Edu.	High-Edu.
$u_t$	-0.011*** (0.002)	-0.014*** (0.002)	-0.008*** (0.003)
$\ln(\Lambda_b + 1)$	0.094*** (0.008)	0.100*** (0.010)	0.105*** (0.014)
$exp \ln(\Lambda_b + 1)$	0.001 (0.001)	-0.002 (0.001)	0.004*** (0.001)
Observations	33386	20885	11722
$R^2$	0.645	0.565	0.675

Standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

low and highly educated workers, but it rejects it when the high skilled are analyzed separately. The regression results suggests that the scale parameter decreases with experience and age for the low-educated workers whereas it increases for the highly educated. A one standard deviation increase in experience (about 6.5 years) raises  $\sigma$  by about 25% for the high educated.

Our model assigns a clear role to lay-offs. After a lay-off, the job search process has to start all over again from the lowest rung of the job ladder. Suppose that this were not true, but that the returns to search accumulate during the career, irrespective of the lay-off of a worker. Then, the estimated effects of the log number of offers during this employment cycle,  $\ln(\Lambda_b + 1)$ , is just a proxy for the omitted variable, the log number of offers accumulated during the whole career,  $\ln(\Lambda_{Tb} + 1)$ . The latter is correlated to  $\ln(\Lambda_b + 1)$  by construction, since both variables are equal during the first cycle. To test this, we run a regression including both variables, see Table 7. For both low- and high-educated workers, the number of offers during the current cycle turns out to be the most important and highly significant, although for the low-educated the number of offers during the whole career matters, as well. This might be explained by misclassification of the reason for separation. If some quits are classified as lay-offs, we would expect  $\ln(\Lambda_{Tb} + 1)$  to be positive, even if the OJS model was correct.

In Table 8 we include —next to  $\ln(\Lambda_b + 1)$ — the variable  $\ln(\lambda b + 1)$ , where  $\lambda$  measures the average number of offers per unit of calendar time and where

Table 7: Robustness test for OJS

	All Data	Low-Edu.	High-Edu.
$u_t$	-0.012*** (0.002)	-0.015*** (0.002)	-0.008*** (0.003)
$\ln(\Lambda_b + 1)$	0.094*** (0.008)	0.065*** (0.010)	0.144*** (0.015)
$\ln(\Lambda_{Tb} + 1)$	0.024* (0.013)	0.050*** (0.016)	-0.020 (0.023)
Observations	33386	20885	11722
$R^2$	0.645	0.565	0.675

Standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ 

$b$  measures calendar time in the current employment cycle. Hence,  $\Lambda_b$  and  $\lambda b$  measure the time elapsed during the same interval, but using different clocks, labour market time in the case of  $\Lambda_b$  and calendar time in the case of  $\lambda b$ . By entering  $\ln(\Lambda_b + 1)$  and  $\ln(\lambda b + 1)$  simultaneously, we can test whether labour market time is indeed the relevant variable, as our model predicts. In the same spirit, we run regressions where we only include the calendar time of an employment run,  $\ln(\lambda b + 1)$ , to test whether the variation in the length of employment cycles due to the random arrival of lay-offs alone can identify the effect of OJS on wages.

Columns 1-3 report the results with both calendar- and labour market time. Labour market time outperforms calendar time. The coefficients on calendar time have a negative sign, though it is only significant when combining low- and highly educated workers and the standard errors are very large. Columns 4-6 report the results with only calendar time. This yields similar estimates for  $\sigma$  as when one only includes labour market time; the coefficient on  $\ln(\lambda b + 1)$  varies between 0.10 and 0.15. Hence, both sources of variation, the random arrival of lay-offs and the business cycle fluctuations in the job-offer arrival rate, can separately identify the effect of OJS - and they both yield estimates for  $\sigma$  of similar magnitude.

In Table 9 we estimate the variance of match quality separately for the rural and urban samples. The variance is greater for those working in the city, in particular for highly educated workers. This is consistent with models

Table 8: Robustness test for OJS

	Calendar and labour market time			Only calendar time		
	All Data	Low-Edu.	High-Edu.	All Data	Low-Edu.	High-Edu.
$u_t$	-0.010*** (0.002)	-0.013*** (0.002)	-0.007** (0.003)	-0.013*** (0.002)	-0.015*** (0.002)	-0.010*** (0.003)
$\ln(\Lambda_b + 1)$	0.265*** (0.092)	0.203* (0.111)	0.254 (0.172)			
$\ln(\lambda b + 1)$	-0.161* (0.092)	-0.117 (0.111)	-0.120 (0.173)	0.104*** (0.005)	0.086*** (0.006)	0.134*** (0.010)
Observations	33386	20885	11722	33386	20885	11722
$R^2$	0.645	0.564	0.675	0.644	0.564	0.675

Standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ 

Table 9: Estimation for subgroups

	Low-Edu.		High-Edu.	
	Rural	Urban	Rural	Urban
$u_t$	-0.010** (0.004)	-0.015*** (0.002)	0.003 (0.009)	-0.010*** (0.004)
$\ln(\Lambda_b + 1)$	0.073*** (0.016)	0.074*** (0.008)	0.020 (0.039)	0.138*** (0.013)
Observations	3786	13589	1124	8036
$R^2$	0.584	0.571	0.705	0.666

Standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ 

with returns to scale in job search where search intensive activities are located in cities (see [Gautier and Teulings \(2009\)](#) and [Elliott \(2014\)](#)).

As discussed in Section 2.1 (see statement (3) and (4)), the expected productivity of a match depends on whether it ends with a quit or a lay-off. The difference in intercept should be equal to  $\sigma$ . As a further test we run the regressions of wages on labour market history for both groups separately. Table 10 reports the results, allowing the intercept and slope to be different for workers subsequently moving into unemployment compared to those that move to another job. As suggested by the theory, the difference in the coefficient for  $\ln(\Lambda_b + 1)$  is small. The difference in means captured by the dummy for quits is close to the estimate of  $\sigma$ .

Table 10: Regressions using the reason for separation

	All Data	Low-Edu.	High-Edu.
$u_t$	-0.011*** (0.002)	-0.015*** (0.003)	-0.008** (0.004)
$\ln(\Lambda_b + 1)1(Fired)$	0.079*** (0.009)	0.075*** (0.010)	0.097*** (0.019)
$\ln(\Lambda_b + 1)1(Quit)$	0.104*** (0.009)	0.080*** (0.012)	0.121*** (0.016)
Quit	-0.125*** (0.015)	-0.101*** (0.017)	-0.114*** (0.034)
Observations	18990	11840	6730
$R^2$	0.707	0.607	0.747

Standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ 

Conditional on the number of offers, the variance of wages is constant for the Gumbel distribution. However, the actual number of offers conditional on its expectation is random, following a log Poisson distribution. The variance of the log number of offers is approximately  $(\Lambda_b + 1)^{-1}$  (this is discussed in relation to equation (9)). To test this prediction, we regress the squared residuals of the wage regression on individual controls and  $(\Lambda_b + 1)^{-1}$ . Since this squaring of the residuals yields a heavily right tailed distribution, we exclude the largest 5% of the squared residuals to reduce the impact of outliers. The results appear in Table 11. We find a positive coefficient, as suggested by the theory. The variance of wages is highest at the beginning of an employment cycle, since the sensitivity of the worker's wage to an additional wage offer is higher in the beginning of a new employment cycle. If the job offers are Poisson distributed we would expect the coefficient to be equal to  $\sigma^2 \cong 0.01 - 0.02$ . The estimated coefficient is about twice as large. This indicates that the distribution might be thinner tailed than the Gumbel distribution, since then the variance of wages conditional on the number of offers  $n$  is decreasing in  $n$  (see equation 4).

Table 11: Regressions for the variance of wages

	All Data	Low-Edu.	High-Edu.
$(\Lambda_b + 1)^{-1}$	0.027*** (0.003)	0.017*** (0.003)	0.043*** (0.005)
Observations	31717	19841	11136
$R^2$	0.026	0.021	0.043

Standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

### 3.5 Applications of the OJS model

We can use this framework to analyse the contribution of job search to wage dispersion. The variance of log wages can be decomposed into three orthogonal components: (i) observed and unobserved general human capital, (ii) random shocks, and (ii) match quality:

$$\text{Var} [\bar{w}_{it}] = \text{Var} [\beta_i + \beta' X_{it}] + \text{Var} [\varepsilon_{it}] + \text{Var} [w_{ia}].$$

The latter term can be further decomposed in three orthogonal terms: (i) the length of the current employment cycle until the end of the current job  $\Lambda_b$  (i.e. variation due to the random arrival of lay-off shocks), (ii) the number of job offers received, conditional on  $\Lambda_b$ , and (iii) conditional on the number of offers received, the distribution the maximum wage offer. For the case of the Gumbel distribution, we obtain a particularly simple formula:

$$\begin{aligned} \text{Var} [w_{ia}] &= \text{Var} [E [w_{ia} | \Lambda_b]] + E [\text{Var} [\ln n | \Lambda_b]] \sigma^2 + \text{Var} [w_{ia} | n] \quad (9) \\ &\cong \left( \text{Var} [\ln(\Lambda_b + 1)] + E [\Lambda_b^{-1}] + \frac{\pi^2}{6} \right) \sigma^2, \end{aligned}$$

where  $\pi = 3.14$ . We apply the first-order approximation for the variance of the log Poisson distribution.<sup>17</sup> Furthermore, we use  $\text{Var} [w_{ia} | n] = \frac{\pi^2}{6} \sigma^2$ ; this expression does not depend on  $n$  for the Gumbel distribution (see equation (4)). Estimates for the first two terms can be obtained from the data. The

<sup>17</sup>

$$\text{Var} [\ln n | \Lambda_b + 1] \cong \left( \frac{d \ln E [n]}{d E [n]} \right)^2 \text{Var} [n | \Lambda_b + 1] = \frac{\Lambda_b}{(\Lambda_b + 1)^2} = \Lambda_b^{-1} + O(\Lambda_b^{-2}),$$

since  $\text{Var} [n | \Lambda_b + 1] = \Lambda_b$  and  $E[n] = \Lambda_b + 1$ .

value of  $\sigma$  is derived from the estimation of equation (5). We can use our estimate of  $\sigma$  to calculate the contribution of OJS to the wage dispersion (see equation (9)). We apply a benchmark estimate of the coefficient  $\sigma$  of 0.10 and use the variance of  $\ln(\Lambda_b + 1)$  from Table 1. We obtain the following result:

Table 12: Variance decomposition

Var[E [ln $n$ ]]	Var [ln( $\Lambda_b + 1$ )] $\sigma^2$	0.0075
Var[ln $n$  E [ln $n$ ]]	E [ $\Lambda_b^{-1}$ ] $\sigma^2$	0.0016
Var[ $w_{ia}$   $n$ ]	$\frac{\pi^2}{6} \sigma^2$	0.0176
Var[ $w_{ia}$ ]	$\left( \text{Var} [\ln(\Lambda_b + 1)] + \text{E} [\Lambda_b^{-1}] + \frac{\pi^2}{6} \right) \sigma^2$	0.0266
Var[ $\bar{w}_{it}$ ]		0.2970
Share	Var[ $w_{ia}$ ] / Var[ $\bar{w}_{it}$ ]	9%

The model can also be applied for the calculation of the expected wage loss after lay-off by comparing the wage in the job from which the worker is fired and in the first job after the lay-off. This is equivalent to comparing the expected wage in the last job of the current employment cycle (for which  $\text{E}[n] = \Lambda_b + 1$ ) with the expected log wage in the first job of the new employment cycle (for which  $n = 1$ ). Using the average value of  $\delta_t$  and  $\lambda_t$  the expected loss in log wages can be calculated as

$$\begin{aligned}
 & \text{E} [w_{ia} | \Lambda_b] - \text{E} [w_{ia} | \Lambda_b = 0] \\
 &= \int_0^\infty \text{Pr}(t) \sigma \ln(\lambda t + 1) dt = -\sigma \exp(\delta/\lambda) \text{Ei}(-\delta/\lambda) \quad (10)
 \end{aligned}$$

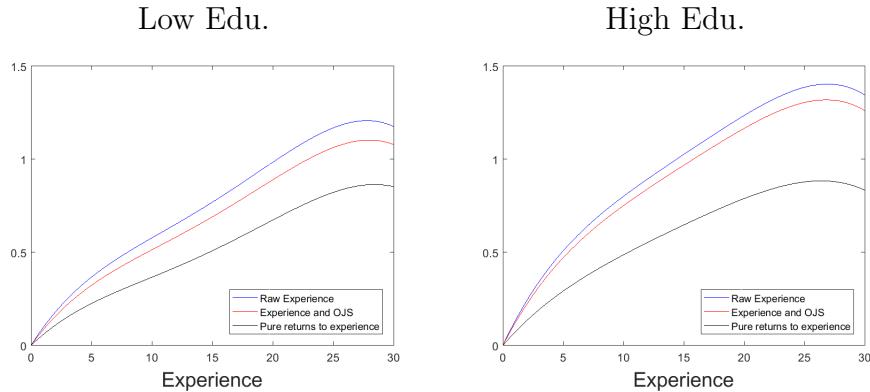
since  $\text{Pr}(t) = \delta \exp[-\delta t]$ . The result from Table 10 provides an estimate for the relation between the experience level of the worker and the wage loss associated with separation. The result suggests that doubling the labour market history increases this wage loss by 8%. Using the steady-state distribution of experience, we derived a simple expression (equation (10)) for the average wage loss in terms of just two parameters,  $\lambda/\delta$  and  $\sigma$ .  $\lambda/\delta$  measures the job-offer arrival rate relative to the rate of separation. If this ratio is high, workers receive more offers on average before they exogenously separate. The higher is  $\lambda/\delta$ , the better are the outstanding matches compared to the average match of a worker coming out of unemployment.  $\sigma$  measures the

scale of the distribution. For a given drop in match quality, the loss in wages increases in  $\sigma$ . Using  $\sigma = 8\%$  and our estimate of  $\lambda/\delta \cong 4$ , yields an average wage loss of about 11%. In addition, our estimates suggest that the offer distribution is fat-tailed, which means that the earnings losses are persistent. Our estimate of 11% is smaller than the empirical estimates of earnings losses following mass displacement (Jacobson et al. (1993) and Davis and Wachter (2011)), but these studies restrict analysis to high tenured workers who on average have a higher match quality. Davis and Wachter (2011) emphasize the importance of the labour market conditions at the time of separation. The job-arrival rates vary roughly by a factor of two: the simple model with OJS therefore results in an earnings loss that is twice as persistent in recessions.

The increase in workers wages over the life cycle has several potential explanations. We seek to decompose the increase into three components: (i) the accumulation of general human capital; (ii) a pure tenure profile in wages; and (iii) the selection into better matches due to OJS. We do this by (i) calculating an “total” increase in wages with experience without controlling for tenure or match quality (ii) estimate increase in wages with experience controlling for the returns to tenure (iii) estimate increase in wages with experience controlling for the returns to tenure and OJS. If we change the order of the decomposition the result remains the same. We use our methodology for estimating the return to OJS and making this decomposition. First, we obtain the total experience profile by running a wage regression with a fourth-order polynomial in experience with the same controls as in all of our previous regressions, but omitting  $\ln(\Lambda_b + 1)$  and the polynomial in tenure. We use the estimated coefficient on the polynomial in experience to generate a predicted experience profile. This gives the total return to experience. The estimated tenure profile would be upward biased if we do not control for match quality due to survival bias. In order to quantify the contribution of tenure to the experience profile, we need an unbiased estimate of the tenure profile. First we derive an unbiased tenure profile, which we then use to correct for the effect of tenure on the experience profile. Regressing wages on match quality  $\ln(\Lambda_b + 1)$ , controls, and correcting for tenure and experience up to a fourth-order polynomial yields us an estimate of the pure tenure profile. For each observation we subtract from the wage the predicted contribution of tenure. We regress the tenure-corrected wage on a fourth-order polynomial in experience and controls. The predicted experience profile includes the returns to OJS and the pure returns to experience. The gap between this experience profile and the total experience profile is the contribution of

tenure to the experience profile. Finally, we get the pure experience profile by regressing wages on fourth-order polynomials in experience and tenure including  $\ln(\Lambda_b + 1)$ . The estimated coefficients on the polynomial yields the pure experience profile. These estimates appear in Figure 5. The return to OJS explains a large part of the total return to experience, (some 30%), and results in a much flatter experience profile. The contribution of tenure to the total return to experience is small.

Figure 5: Experience Profile with and without controlling for OJS

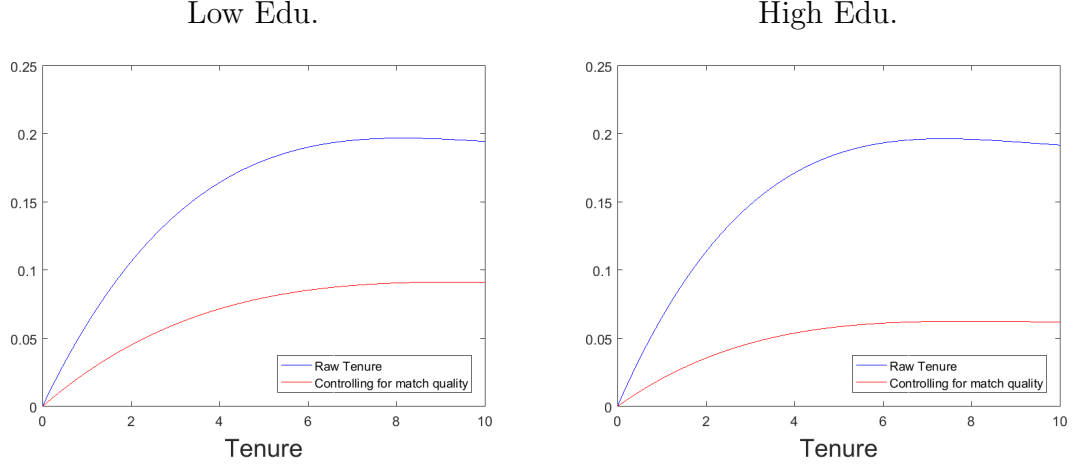


We perform a similar decomposition for the tenure profile. First, we obtain the total return to tenure by running a wage regression with the standard controls and fourth-order polynomials in experience and tenure but omitting  $\ln(\Lambda_b + 1)$ . Next, we run the same regression, but include  $\ln(\Lambda_b + 1)$ . The tenure profiles derived from both regressions appear in Figure 6. The results suggest that most of the raw tenure profile is due to survival bias. This result explains why we find that  $\Lambda_a/\Lambda_b$  is uniformly distributed. The return to job-specific experience and the true tenure profile in wages are apparently not that important.

Job-offer arrival rates differ between individuals. Some job seekers receive offers frequently, while others have to wait a long time. As a robustness check, we investigate how our analysis should be adapted when the job-offer arrival rate depends on time-invariant observable human capital of the job



Figure 6: Tenure Profile with and without controlling for OJS



seeker, denoted  $X_i$ .<sup>18</sup> We assume that the arrival rate follows the well-known proportional hazard model:  $\lambda_t \exp(\theta' \tilde{X}_i)$ ,  $\tilde{X}_i$  is the deviation of  $X_i$  from its mean;  $\exp(\theta' \tilde{X}_i)$  is the baseline hazard rate, and  $\theta$  is a parameter vector. Compared to equation (5), we should replace  $\Lambda_b$  by  $\Lambda_b \exp(\theta' \tilde{X}_i)$ . Again, focusing on the Gumbel distribution for the sake of convenience, we obtain

$$\begin{aligned}
 E[w_{ia}|\Lambda_b] &\cong E[w_{ia}|E[n|\Lambda_b]] = \mu + \sigma\gamma + \sigma \ln \left[ \Lambda_b \exp(\theta' \tilde{X}_i) + 1 \right] \\
 &= \mu + \sigma\gamma + \sigma \ln \left[ \Lambda_b + \exp(-\theta' \tilde{X}_i) \right] + \sigma\theta' \tilde{X}_i \\
 &\cong \mu + \sigma\gamma + \sigma \ln(\Lambda_b + 1) + \sigma\theta' \tilde{X}_i + O(\text{Var}[X_i]).
 \end{aligned}$$

Differences in the job-offer arrival rate between individuals are therefore absorbed in the fixed effects up to a term of order  $\text{Var}[X_i]$ . As long as the coefficient of variation of  $X_i$  is small relative to  $\Lambda_b$ , variation in the hazard rate is absorbed in the term  $\beta_i$  in equation (8).

<sup>18</sup>In the current specification,  $X_i$  can include the fixed worker effect  $\beta_i$ ; it cannot include experience, since then  $\lambda_t$  would depend on time not only due to the business cycle but also due to the return to experience.

## 4 Conclusion

No single structural model will be able to fully explain the empirical pattern of job-to-job transitions and wage-dynamics. Too many mechanisms play a role for all of them to be captured in a single structural framework. From that perspective, the results presented in this paper are surprising. We posited the simple common framework used in [Burdett and Mortensen \(1998\)](#), [Bontemps et al. \(2000\)](#), and [Shimer \(2006\)](#), with common job offer arrival and job destruction rates where workers accept any offer that pays more than their current job. This simple structure has many testable implications both for wages and job durations. Reiterating the main implications: the selection process proceeds at the same speed during subsequent employment cycles, a lay-off restarts the selection process, jobs pay lower wages when ending in a quit rather than a lay-off, and the duration of the current job is uniformly distributed over the length current employment cycle. Apart from some minor deviations, these implications have all been confirmed by the data. In particular the confirmation of the final implication -the uniform distribution of the job duration over the current employment cycle- came as a surprise. One would expect the quit rate to go down over the duration of a job by workers acquiring firm specific human capital or by tenure premiums paid by firms as a retention bonus, see [Burdett and Coles \(2003\)](#), implying that the current job is more likely to have started in the beginning rather than the end of the current employment cycle.

Our empirical inference is based on a rigorous derivation of the model's implications for the distribution of job durations over the course of an employment cycle and the implications for the expected quality of a job at different points during the employment cycle. We have shown that the length of the employment cycle at the moment that a job ends is a sufficient statistic for the expected quality of a job. Surprisingly, the starting date of a job does not provide any additional information. We show that one has to correct for the difference between jobs ending by a quit rather than a lay-off, since an ending by a quit reveals that still better jobs were available, which is a negative sign on the quality of the current job. Our methodology allow us to estimate the contribution of OJS to the dispersion of wages. Search frictions explain 9% of wage dispersion.

Our method allows inference about the shape of the wage offer distribution. Though many distribution are consistent with the data, our results rule out the uniform distribution, and *a fortiori* distributions where the density

reaches a maximum at its upper support. This is important as it implies that search frictions cannot be fully ascribed to sorting models in which the optimal assignment is an interior point of the matching set, as [Shimer and Smith \(2000\)](#) and [Gautier et al. \(2010\)](#), since these models yield a peak in the density at its upper support. Hence, sorting can only offer a partial explanation for the wide dispersion of the wage offer distribution.

The assumptions applied in this paper are consistent with the models by [Burdett and Mortensen \(1998\)](#), [Bontemps et al. \(2000\)](#), [Pissarides \(1994\)](#) and [Shimer \(2006\)](#), but not with the sequential auctioning model by [Postel-Vinay and Robin \(2002\)](#). We assume that wages within a job increase only due to general human capital, not as a result of receiving outside offers that drive up the wage as in [Postel-Vinay and Robin \(2002\)](#). However, since the transition dynamics of both models are the same (workers accept any job offer that is "better" than their current job), our methodology can be easily adapted to encompass this wage determination mechanism and to test its implication, as we do in accompanying paper, see [Gottfries and Teulings \(2017a\)](#). Similarly, our methodology can be adapted to investigate how wages respond to cyclical fluctuations in the job offer arrival rate, as we do in another paper, see [Gottfries and Teulings \(2017b\)](#).

Given the robustness of the support for the basic framework, OJS is of first order importance for any model of the labour market. There might be learning as in [Moscarini \(2005\)](#), there might be adverse selection as in [Gibbons and Katz \(1991\)](#), but as a first order approximation one does not make a big mistake by focusing on the process of gradual selection into better matches over the course of an employment cycle and the restart of this process after a lay-off. A macro model for the labour market that accounts for this process is likely to have the first order phenomena right.

## References

- ARNOLD, B. C., N. BALAKRISHNAN, AND H. NAGARAJA (2008): “A First Course in Order Statistics (Classics in Applied Mathematics),” .
- BAGGER, J., F. FONTAINE, F. POSTEL-VINAY, AND J.-M. ROBIN (2014): “Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics,” *American Economic Review*, 104, 1551–96.
- BARLEVY, G. (2008): “Identification of Search Models using Record Statistics,” *The Review of Economic Studies*, 75, 29–64.
- BONTEMPS, C., J.-M. ROBIN, AND G. V. DEN BERG (2000): “Equilibrium Search with Continuous Productivity Dispersion: Theory and Non-parametric Estimation,” *International Economic Review*, 41, pp. 305–358.
- BURDETT, K. AND M. COLES (2003): “Equilibrium Wage-Tenure Contracts,” *Econometrica*, 71, 1377–1404.
- BURDETT, K. AND D. T. MORTENSEN (1998): “Wage differentials, employer size, and unemployment,” *International Economic Review*, 257–273.
- DAVIS, S. J. AND T. V. WACHTER (2011): “Recessions and the Costs of Job Loss,” *Brookings Papers on Economic Activity*, 43, 1–72.
- DE HAAN, L. AND A. FERREIRA (2007): *Extreme value theory: an introduction*, Springer Science & Business Media.
- DREW, JULIA A. RIVERA, S. F. AND J. R. WARREN (2014): “Making full use of the longitudinal design of the Current Population Survey: Methods for linking records across 16 months,” *Journal of Economic and Social Measurement*, 39, 121–144.
- DUSTMANN, C. AND C. MEGHIR (2005): “Wages, Experience and Seniority,” *The Review of Economic Studies*, 72, 77.
- ELLIOTT, M. (2014): “Heterogeneities and the Fragility of Labor Markets,” Available at SSRN 2445753.

- FREDRIKSSON, P., L. HENSVIK, AND O. NORDSTROM SKANS (2015): “Mismatch of Talent: Evidence on Match Quality, Entry Wages, and Job Mobility,” *IZA Discussion Paper No. 9585*.
- GAUTIER, P. A. AND C. N. TEULINGS (2009): “Search and the city,” *Regional Science and Urban Economics*, 39, 251–265.
- (2015): “Sorting and the output loss due to search frictions,” *Journal of the European Economic Association*.
- GAUTIER, P. A., C. N. TEULINGS, AND A. VAN VUUREN (2010): “On-the-Job Search, Mismatch and Efficiency,” *The Review of Economic Studies*, 77, 245–272.
- GIBBONS, R. AND L. F. KATZ (1991): “Layoffs and lemons,” *Journal of Labor Economics*, 9, 351–380.
- GOTTFRIES, A. AND C. N. TEULINGS (2017a): “How are wages set? An empirical test with on-the-job-search,” .
- (2017b): “Wage posting, nominal rigidity, and cyclical inefficiencies,” .
- GUVENEN, F., B. KURUSCU, S. TANAKA, AND D. WICZER (2015): “Multidimensional Skill Mismatch,” NBER Working Papers 21376, National Bureau of Economic Research, Inc.
- HAGEDORN, M. AND I. MANOVSKII (2013): “Job Selection and Wages over the Business Cycle,” *The American Economic Review*, 103, pp. 771–803.
- JACOBSON, L. S., R. J. LALONDE, AND D. G. SULLIVAN (1993): “Earnings Losses of Displaced Workers,” *American Economic Review*, 83, 685–709.
- MANOVSKII, I. AND M. HAGEDORN (2010): “Search Frictions and Wage Dispersion,” 2011 Meeting Papers 1195, Mineo.
- MOSCARINI, G. (2005): “Job Matching and the Wage Distribution,” *Econometrica*, 73, 481–516.
- PISSARIDES, C. A. (1994): “Search Unemployment with On-the-job Search,” *The Review of Economic Studies*, 61, 457–475.

- POSTEL-VINAY, F. AND J. LISE (2015): “Multidimensional Skills, Sorting, and Human Capital Accumulation,” 2015 Meeting Papers 386, Society for Economic Dynamics.
- POSTEL-VINAY, F. AND J.-M. ROBIN (2002): “Equilibrium Wage Dispersion with Worker and Employer Heterogeneity,” *Econometrica*, 70, pp. 2295–2350.
- SHIMER, R. (2006): “On-the-job search and strategic bargaining,” *European Economic Review*, 50, 811–830.
- SHIMER, R. AND L. SMITH (2000): “Assortative Matching and Search,” *Econometrica*, 68, pp. 343–369.
- TJADEN, V. AND F. WELLSCHMIED (2014): “Quantifying the Contribution of Search to Wage Inequality,” *American Economic Journal: Macroeconomics*, 6, 134–161.
- WOLPIN, K. I. (1992): “The Determinants of Black-White Differences in Early Employment Careers: Search, Layoffs, Quits, and Endogenous Wage Growth,” *Journal of Political Economy*, 100, pp. 535–560.

## A Proof of Proposition 1

1. The expected number of job offers  $n$  in the time interval  $[0, b]$ :

Define  $k$  to be the number of job offers in the time interval  $(0, b)$ .

- (a) Consider the case of a job ending by a lay-off. By definition, the worker receives one offer at time 0, since at that time he transits from unemployment to employment. The probability of an offer exactly at time  $b$  has measure zero. Hence,  $k = n - 1$ . The distribution of  $k$  conditional termination via a lay-off at time  $b$  and  $\Lambda_b$  is

$$\begin{aligned} \Pr(k|\Lambda_b, \text{lay-off at } b) &= \frac{\Pr(k, \text{lay-off at } b|\Lambda_b)}{\Pr(\text{lay-off at } b|\Lambda_b)} \\ &= \Pr(k|\Lambda_b) \frac{\Pr(\text{lay-off at } b|k, \Lambda_b)}{\Pr(\text{lay-off at } b|\Lambda_b)} = \Pr(k|\Lambda_b). \end{aligned}$$

where the final equality uses that the separation shock is independent of the number of previous offers.  $k$  follows a Poisson distribution with parameter  $\Lambda_b$ . Since the expectation of this distribution is  $\Lambda_b$ , we have

$$E[n] = 1 + \Lambda_b.$$

- (b) Consider the case of a job ending by a quit. Again, by definition, the worker must have receive an offer at time 0. However, in this case the worker must also have received an offer at time  $b$ , for otherwise he could not have quited for a better job. Hence,  $k = n - 2$ . Conditional on the arrival of an offer at time  $b$ , the probability that this offer leads to a quit is  $n^{-1}$ , since for this offer to lead to a quit, it must be the highest of  $n$  offers and since all i.i.d. draws in a sequence of  $n$  are equally likely to be the maximum. The distribution of  $k$  conditional termination via a lay-off at time  $b$  and  $\Lambda_b$  is

$$\begin{aligned} \Pr(k|\Lambda_b, \text{quit at } b) &= \frac{\Pr(k, \text{quit at } b|\Lambda_b, \text{offer at } b)}{\Pr(\text{quit at } b|\Lambda_b, \text{offer at } b)} \\ &= \Pr(k|\Lambda_b) \frac{\Pr(\text{quit at } b|k, \Lambda_b, \text{offer at } b)}{\Pr(\text{quit at } b|\Lambda_b, \text{offer at } b)} \\ &= \frac{\Lambda_b^k e^{-\Lambda_b}}{k! (k+2)} \Pr(\text{quit at } b|\Lambda_b, \text{offer at } b)^{-1}. \end{aligned}$$

using  $\Pr(k|\Lambda_b, \text{offer at } b) = \Pr(k|\Lambda_b)$ . Hence

$$\begin{aligned} \mathbb{E}[k + 2|\Lambda_b, \text{quit at } b] &= \left( \sum_{k=0}^{\infty} \frac{\Lambda_b^k}{k!(k+2)} \right)^{-1} \sum_{k=0}^{\infty} \frac{\Lambda_b^k}{k!} = \frac{\Lambda_b^2}{\Lambda_b - (1 - e^{-\Lambda_b})} \\ &= \Lambda_b + 1 + O(\Lambda_b^{-1}). \end{aligned}$$

2.  $\Lambda_a/\Lambda_b$  is uniformly distributed on the unit interval  $[0, 1]$ :

The probability that the length  $b$  of a worker's employment cycle exceeds  $t$  and that his job has a rank less than  $F$  satisfies

$$\Pr(\underline{F} \leq F, b > t|t) = F \times \exp[-\Lambda_t(1 - F) - \Delta_t], \quad (11)$$

The first factor  $F$  is the probability that the first offer of the employment cycle at time 0 is less than  $F$ , while the second factor is the probability that now offer higher than  $F$  has come in during the time interval  $(0, t)$ . Conditional on the starting date of the job  $a$  and its rank  $F$ , the end date  $b$  satisfies

$$\begin{aligned} \Pr(b|F, a) &= [\lambda_b(1 - F) + \delta_b] \times \\ &\quad \exp[-(\Lambda_b - \Lambda_a)(1 - F) - (\Delta_b - \Delta_a)]. \end{aligned} \quad (12)$$

The joint density of  $a, b, F$  is the product of four probabilities: (i) the probability that there is an offer at  $a$ , which is  $\lambda_a$ ; (ii) the probability that this offer has rank  $F$ , which is unity since  $F$  is uniformly distributed; (iii) the probability that there has been no prior offer greater than  $F$  conditional on  $a$ ; and (iv) the probability that this job ends at  $b$  conditional on  $F$  and  $a$ . Equation (11) for  $t = a$  gives the probability (iii). Equation (12) gives probability (iv). Hence

$$\begin{aligned} \Pr(F, b, a) &= \Pr(\text{offer at } a) \Pr(F|\text{offer at } a) \Pr(\underline{F} < F, b > a|a) \Pr(b|F, a) \\ &= \lambda_a F \exp[-\Lambda_b(1 - F) - \Delta_b] [\delta_b + \lambda_b(1 - F)]. \end{aligned}$$

Integrating over  $F$  and dividing by the probability of  $b$  gives

$$\Pr(a|b) = \lambda_a C_3(b)^{-1}.$$

Since  $a \in [0, b]$ ,  $C_3(b) = \int_0^b \lambda_a da = \Lambda_b$ . The distribution of  $a$  is therefore

$$\Pr(\underline{a} \leq a|b) = \frac{\Lambda_a}{\Lambda_b}.$$

$\Lambda_a/\Lambda_b$  is therefore uniformly distributed.



3. Expected rank for jobs ending in a lay-off:

$$\begin{aligned}\Pr(F|\Lambda_b, \text{lay-off at } b) &= \frac{\Pr(F, \text{lay-off at } b|\Lambda_b)}{\Pr(\text{lay-off at } b|\Lambda_b)} \\ &= \Pr(F|\Lambda_b) \frac{\Pr(\text{lay-off at } b|\Lambda_b, F)}{\Pr(\text{lay-off at } b|\Lambda_b)} = \Pr(F|\Lambda_b),\end{aligned}$$

where we use the fact that the probability of a lay-off at  $b$  is independent of the value of  $F$  in the last equality. The cumulative distribution function of  $F$  conditional on the initial offer at time 0 and  $\Lambda_b$  is

$$\Pr(\underline{F} \geq F|\Lambda_b) = F \exp[-\Lambda_b(1 - F)].$$

Hence

$$\Pr(F|\Lambda_b) = (\Lambda_b F + 1) \exp[-\Lambda_b(1 - F)]. \quad (13)$$

Using this, the expectation of  $F$  can be written as

$$\begin{aligned}\mathbb{E}(F|\Lambda_b) &= \int_0^1 F (\Lambda_b F + 1) \exp[-\Lambda_b(1 - F)] dF \\ &= 1 - \Lambda_b^{-1} + \Lambda_b^{-2} [1 - \exp(-\Lambda_b)] = 1 - \Lambda_b^{-1} + O(\Lambda_b^{-2}).\end{aligned}$$

4. Expected rank for jobs ending in a quit:

$$\begin{aligned}\Pr(F|\Lambda_b, \text{quit at } b) &= \frac{\Pr(F, \text{quit at } b|\Lambda_b)}{\Pr(\text{quit at } b|\Lambda_b)} \\ &= \frac{\Pr(F|\Lambda_b) \Pr(\text{quit at } b|\Lambda_b, F)}{\Pr(\text{quit at } b|\Lambda_b)}\end{aligned}$$

By equation (13) and since  $\Pr(\text{quit at } b|\Lambda_b, F) = \lambda_b(1 - F)$ , the conditional distribution is therefore

$$\Pr(F|\Lambda_b, \text{quit at } b) = C_4^{-1} (1 - F) (\Lambda_b F + 1) \exp[-\Lambda_b(1 - F)],$$

where

$$\begin{aligned}C_4 &= \left[ \int_0^1 (1 - F) (\Lambda_b F + 1) \exp[-\Lambda_b(1 - F)] dF \right]^{-1} \\ &= \Lambda_b^{-1} (1 - \Lambda_b^{-1} [1 - \exp(-\Lambda_b)]).\end{aligned}$$

The expected match quality is

$$\begin{aligned}\mathbb{E}(F|\Lambda_b, \text{quit at } b) &= C_4^{-1} \int_0^1 F (1 - F) (\Lambda_b F + 1) \exp[-\Lambda_b(1 - F)] dF \\ &= 1 - 2\Lambda_b^{-1} + O(\Lambda_b^{-2}).\end{aligned}$$

## B Proof of proposition 2

The distribution of the number of offers given a lay off is

$$\Pr(n|\Lambda_b, \text{Lay-off}) = \frac{\Lambda_b^{n-1} e^{-\Lambda_b}}{(n-1)!}.$$

The expectation given  $\Lambda_b$  can therefore be written as

$$\mathbb{E}[w|\Lambda_b, \text{Lay-off}] = \sum_{n=1}^{\infty} \frac{\Lambda_b^{n-1} e^{-\Lambda_b}}{(n-1)!} w_n.$$

where  $w_n$  the expected maximum from  $n$  draws. First we prove that the function  $\mathbb{E}[w|\Lambda_b, \text{Lay-off}]$  characterises the sequence  $w_n$ . This equation holds identically for all  $\Lambda_b$ . Hence, its derivative with respect to  $\Lambda_b$  should also apply. Multiplying both sides by  $e^{\Lambda_b}$  and taking the  $k^{\text{th}}$  derivative yields

$$e^{\Lambda_b} \left[ \sum_{i=0}^k \frac{k!}{(k-i)!i!} \frac{\partial^i \mathbb{E}[w|\Lambda_b, \text{Lay-off}]}{\partial \Lambda_b^i} \right] = \sum_{n=k+1}^{\infty} \frac{\Lambda_b^{n-k-1}}{(n-k-1)!} w_n,$$

where we adopt the convention that the  $(i=0)^{\text{th}}$  derivative of a function is the function itself. This expression holds for all  $\Lambda_b$ . Evaluating it for  $\Lambda_b = 0$  yields

$$w_{k+1} = \sum_{i=0}^{k+1} \frac{k!}{(k-i)!i!} \frac{\partial^i \mathbb{E}[w|\Lambda_b, \text{Lay-off}]|_{\Lambda_b=0}}{\partial \Lambda_b^i}.$$

The function  $\mathbb{E}[w|\Lambda_b, \text{Lay-off}]$  therefore characterises the sequence  $w_n$ . The sequence of expected maxima  $w_n$  in turn identifies any distribution, see Theorem 6.3.1 in [Arnold, Balakrishnan, and Nagaraja \(2008\)](#).

## C Identification of $\psi$

For the estimation of  $\psi$ , we rely on a steady-state argument, where labour market time runs at a constant rate. Hence, we drop the suffix  $t$  of  $\lambda_t$  and  $\delta_t$ . First, we derive the expected duration of the first job of an employment cycle. The duration of a job of rank  $F$  follows an exponential distribution with parameter  $\delta + \lambda(1-F)$ . Hence, the expected duration of a job conditional

on its rank is  $[\delta + \lambda(1 - F)]^{-1}$ . Since the rank of the first job is a random draw from the uniform distribution, its expected duration satisfies

$$\mathbb{E}[b|1^{st} \text{ job in emp.cycle}] = \int_0^1 [\delta + \lambda(1 - F)]^{-1} dF = \lambda^{-1} \ln(1 + \lambda/\delta). \quad (14)$$

Next, we derive the expected termination date  $b$  of all subsequent jobs. First, we calculate the joint density among all jobs of the rank  $F$  of the current job, its start date  $a$ , and its termination date  $b$ . This density is comprised of three parts: (i) the fraction  $F \exp[-(\delta + \lambda(1 - F))a]$  of workers remaining at  $a$  with rank less than  $F$ , (ii) the arrival rate  $\lambda$  of an offer at  $a$ , and (iii) the probability  $[\delta + \lambda(1 - F)] \exp[-(\delta + \lambda(1 - F))(b - a)]$  that a match ends at  $b$  conditional on it having started at  $a$ . Hence, this density is proportional to

$$\begin{aligned} \Pr(F, a, b) &\propto F \exp[-(\delta + \lambda(1 - F))a] \times \lambda \times [\delta + \lambda(1 - F)] \exp[-(\delta + \lambda(1 - F))(b - a)] \\ &\propto F \exp[-(\delta + \lambda(1 - F))b] [\delta + \lambda(1 - F)]. \end{aligned}$$

In the second line,  $a$  drops out. We can ignore the job-offer arrival rate  $\lambda$ , since it depends on neither  $F$ , nor  $a$ , nor  $b$ . We integrate this density over the possible start dates  $a \in (0, b)$  to get the joint density of match quality  $F$  and end date  $b$ :

$$\Pr(F, b) = \frac{F \exp[-(\delta + \lambda(1 - F))b] [\delta + \lambda(1 - F)] b}{\frac{\delta + \lambda}{\lambda^2} \ln\left(\frac{\delta + \lambda}{\delta}\right) - \lambda^{-1}},$$

Hence:

$$\mathbb{E}[b|\text{subseq.jobs}] = \int_0^1 \int_0^\infty b \Pr(F, b) db dF = \frac{2}{\lambda} \frac{\lambda/\delta - \ln(1 + \lambda/\delta)}{(1 + \delta/\lambda) \ln(1 + \lambda/\delta) - 1}. \quad (15)$$

We can derive information on  $\mathbb{E}[b|1^{st} \text{ job in emp.cycle}]$  and  $\mathbb{E}[b|\text{subseq.jobs}]$  from the data. This yields a system of two equations, which can be solved for  $\delta$  and  $\lambda$ . The ratio of  $\lambda$  to  $\lambda_u$  provides an estimate for  $\psi$ .

## D Estimation results by employment cycle

Table 13: Estimation for the different employment runs

	All Data	Low-Edu.	High-Edu.
$u_t$	-0.012*** (0.002)	-0.014*** (0.002)	-0.008*** (0.003)
$\ln(\Lambda_b + 1)1(\text{cycle } 1)$	0.121*** (0.011)	0.094*** (0.014)	0.150*** (0.018)
$\ln(\Lambda_b + 1)1(\text{cycle } 2)$	0.120*** (0.008)	0.090*** (0.010)	0.149*** (0.012)
$\ln(\Lambda_b + 1)1(\text{cycle } 3)$	0.123*** (0.007)	0.092*** (0.009)	0.151*** (0.013)
$\ln(\Lambda_b + 1)1(\text{cycle } 4)$	0.111*** (0.008)	0.085*** (0.009)	0.145*** (0.014)
$\ln(\Lambda_b + 1)1(\text{cycle } 5)$	0.097*** (0.009)	0.091*** (0.011)	0.109*** (0.016)
$\ln(\Lambda_b + 1)1(\text{cycle } 6)$	0.113*** (0.010)	0.105*** (0.013)	0.111*** (0.019)
$\ln(\Lambda_b + 1)1(\text{cycle } 7)$	0.095*** (0.012)	0.093*** (0.014)	0.132*** (0.022)
$\ln(\Lambda_b + 1)1(\text{cycle } 8)$	0.075*** (0.012)	0.074*** (0.013)	0.081*** (0.024)
$\ln(\Lambda_b + 1)1(8 < \text{cycle})$	0.059*** (0.010)	0.069*** (0.012)	0.068*** (0.023)
Quit	-0.042*** (0.007)	-0.055*** (0.008)	-0.015 (0.013)
Observations	33386	20885	11722
$R^2$	0.645	0.565	0.676

Standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$