Returns to on-the-job search and the dispersion of wages

Axel Gottfries and Coen Teulings*

August 31, 2016

Abstract

A wide class of models with On-the-Job Search (OJS) predicts that workers gradually select into better-paying jobs, until lay-off occurs, when this selection process starts over from scratch. We develop a simple methodology to test these predictions. Our inference uses two sources of identification to distinguish between returns to experience and the gains from OJS: (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. Using extreme value theory, we can infer the shape of the wage-offer distribution from the effect of the time since the last lay-off on wages.

This methodology is applied to the NLSY 79. We find remarkably strong support for all implications. The offer distribution is Gumbel, which has an unbounded support, which is inconsistent with pure sorting models. The standard deviation of wage offers is 7 to 15% (depending on educational level and urbanisation). OJS accounts for 30% of the experience profile and 9% of total wage dispersion. The average wage loss after lay-off is 11%.

*Gottfries: Faculty of Economics, University of Cambridge, CB3 9DD, United Kingdom, ag558@cam.ac.uk. Teulings: Faculty of Economics, University of Cambridge, CB3 9DD, United Kingdom, cnt23@cam.ac.uk. We would like to thank Jim Albrecht, Jan Eeckhout, Marcus Hagedorn, John Kennan, Philipp Kircher, Moritz Kuhn, Hannes Malmberg, Iourii 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, SaM conferences in Amterdam and Cambridge, NYU macro student lunch and Cambridge University.

Keywords: On-the-job search, Wage dispersion, Job duration
JEL Classification: J31, J63, J64

1 Introduction

Labour market models with on-the-job search (OJS) as developed by Burdett and Mortensen (1998) have become the workhorse model for explaining job-to-job transitions and wage dynamics in macro economics. Since job-to-job flows are an empirically relevant feature of the data (see Nagypal (2006)), a model with OJS is particularly useful. While there are several versions of the model, which differ in particular in their assumptions on wage formation (see for example the model with tenure profiles by Burdett and Coles (2003), the model with sequential auctioning by Postel-Vinay and Robin (2002), or the model with bargaining by Shimer (2006)) the basic mechanics of this class of models are the same. Workers receive job offers as random draws from a stable offer distribution. They accept any offer that improves their expected lifetime utility. In this way they gradually select into better matches, up untill the point where they are laid off. After lay-off, this selection process of increasingly better job offers starts all over again. The models predict an increasing and concave wage-experience profile with workers moving up the job ladder up untill the point of lay-off.

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 transitions and wage growth which can be brought to the data. We find remarkably strong empirical support for a simple version of this model. The returns to OJS are highly stable over the life cycle. These returns explain about 30% of the overall return to labour market experience. The random nature of the returns to OJS accounts for 9% of wage dispersion among male workers. The persistence of the wage loss depends on the arrival rate of job offers and the shape of the offer distribution. If job offers arrive frequently, then workers revert quickly to the top of the job ladder. If the distribution of job offers has a thin upper tail, the gains of selection die out quickly. The methodology laid out in this paper allows the measurement of the job offer arrival rate and the characterization of the shape and the dispersion of the offer distribution.

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 first moves from unemployment to a job and subsequently from one job to another, and ends when the worker gets laid off and becomes unemployed again. This 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. Random variation in the arrival of lay-off shocks that characterize the beginning and the end of subsequent employment cycles can be used to estimate the returns of OJS. Barlevy (2008) shows how results from record theory can be used to characterise the distribution of job offers. The current job offer is viewed as the maximum of the job offers received to date. The record measures the highest draw and is updated every time a higher offer is received. A sequence of jobs in an employment cycle is then a sequence of records. A sequence of wage changes for workers who change jobs compared to the evolution of wages for those who do not change jobs can then be used to characterise the distribution. Our approach uses not only the sequence of records, but also the time between successive records. This gives us more variation to identify the distribution.

Hagedorn and Manovskii (2013) use the interaction between job-to-job transitions and wage dynamics to control for match quality in wage regressions. They apply this methodology to test for history dependence in wages. In a tight labour market, job offers arrive more frequently. Hence, workers move up the job ladder more quickly. Hagedorn and Manovskii (2013) use business cycle fluctuations in labour market tightness to estimate the return to OJS in terms of improvement in match quality.

In order to to disentangle the return to OJS from the general returns to experience this paper combines two sources of variation, individual variation in lay-offs and aggregate variation in job-offer arrival rates related to the business cycle. 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 at which labour market time runs is equal to the job-offer arrival rate for OJS. An attractive feature of our methodology is that it does not depend on the cyclical behaviour of lay off rates.

Our methodology is summarized in three propositions. The first propo-

sition deals with the relation between the expected rank of the actual job in the offer distribution and the observed job tenures within an employment cycle. We show that the end date of the current job is a sufficient statistic for its expected rank. The starting date of that job does not add any useful information. In fact, we show that the model predicts that in the absence of job-specific experience, 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. If job-specific experience were important, the latter result would not hold, as the likelihood of a job transition would decline over the employment cycle (since, by switching jobs, a worker spoils the returns of accumulated job-specific experience). The second and the third proposition are derived from Extreme Value Theory. The second proposition states that there is a one-to-one correspondence between the expected value of the current job as a function of the number of job offers and the shape of the offer distribution. The third propostion discusses the role of Generalized Extreme Value distribution in models of OJS. It deals with the relation between the evolution of the expected rank of the current job over the employment cycle and the shape of the offer distribution.

We apply our methodology to the data from the NLSY 79. We find remarkably strong support for all predictions derived from this model. We show that match quality improves over an employment cycle according to a concave profile. The shape of this profile is the same across subsequent employment cycles in a worker's career, irrespective of whether the employment cycle is the first in his career (starting at the date of labour market entry) or is a subsequent cycle (starting after a lay-off at higher age). We show that both sources of variation, the restart of employment cycles after lay-off and business cycle variation in job offer arrival rates, generate the same estimates of the return to OJS and that ignoring either source of variation yields an underestimation of these returns. Our discussion reveals that the empirical distinction between quits and lay offs as registered in the data matches very well the theoretical distinction between quits and lay-offs in the benchmark OJS model. We show that the starting date of the current job is indeed uniformly distributed over the length of the employment cycle up until the termination date of the current job, suggesting that job specific experience plays a minor role. This conclusion is corroborated by our finding that the tenure profile in wages largely disappears when controlling properly for the return to OJS.

Now that the empirical validity of the model has been established, it

can be applied for empirical inference. We show that the distribution of log wage offers is best characterized by a Gumbel distribution, which has an unbounded upper support and a fat upper 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 upper support of the offer distribution is finite, implying a thin right tail. Hence, sorting cannot be the only source of variation in the offer distribution. The standard deviation of the offer distribution is estimated to be about 10%. We differentiate this estimate by the education level of the workers and between cities on the one hand and the countryside on the other hand. The standard deviation is shown to be one and a half times higher for higher educated workers than for the lower educated and one and a half times higher for the city than for the countryside, where differences between the city and the countryside are more pronounced for higher educated workers than for the lower educated. This provides support for models that explain the existence of cities from returns to scale in job search (including Gautier and Teulings (2009)). These estimates can then be used to derive a number of relevant statistics. As stated before, our estimates suggest that the expected return to OJS explains some 30% of the observed return to experience, and that controlling properly for the return to OJS explains most of the return to tenure. The return to OJS explains some 9% of male wage dispersion, for four reasons: i) on average, the length of the current employment cycle increases with the experience of the worker, ii) random variation occurs in the length of employment cycles due to lay-off shocks; iii) random variation exists in the number of job offers received (it is a Poisson process), and iv) random variation plays a role in the quality of the job offer (see Manovskii and Hagedorn (2010) for an alternative estimate). Finally, our estimates of the wage offer distribution show that the fall in wages in the first job after displacement is about 11%.

This paper relates to a number of recent papers interested in match quality (see, for example, Fredriksson, Hensvik, and Nordstrom Skans (2015), Guvenen, Kuruscu, Tanaka, and Wiczer (2015) and Postel-Vinay and Lise (2015)). Jacobson, LaLonde, and Sullivan (1993) and Davis and Wachter (2011) have shown that the wage losses following mass lay-offs are large and persistent. Bagger, Fontaine, Postel-Vinay, and Robin (2014) and Altonji, Smith, and Vidangos (2013) have provided estimates for the part of the experience profile that comes from moving up the job ladder. This paper assumes that job transitions are efficient, in the sense that workers always move to

better quality jobs when a job offer arrives. Testing this assumption is an interesting topic for future research.

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.

2 The theoretical argument

2.1 Assumptions

Consider a labour market with search frictions and OJS. Worker i's log wage at time t, denoted $\bar{w}_{it}$, satisfies

$$\bar{w}_{it} = \beta_i + \beta' X_{it} + w_{ia} + \varepsilon_{it}, \tag{1}$$

where β_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 ε_{it} is a random variable. The three components $\beta_i + \beta' X_{it}, w_{ia}$ and ε_{it} are mutually uncorrelated. For the moment, we abstract from job-specific work experience; we return to this issue in Section 2.2.

During their labour market career, workers receive job offers at a rate λ_t. When receiving a job offer characterized by w_{ia}, the worker negotiates with the firm on the wage. We do not take a stance on how that bargaining process proceeds, since we are only interested in its outcome, the match-specific component of the log wage, w_{ia}. We assume w_{ia} to be constant over the duration of the job. Let $F(w_{ia})$ denote the distribution function of w_{ia}; $F(\cdot)$ is assumed to be differentiable; alternatively, $F(w_{ia})$ will be referred to as the rank of job offer. We assume that even the least attractive job offer is more attractive than unemployment. Hence, an unemployed job seeker accepts any job offer. The wage-bargaining process is assumed to be efficient: workers accept any job offer that yields a higher output than their current job; they do this because their higher productivity enables the employer to offer a higher wage. Most models of job search have this property. As the labour market history of workers accumulates, they receive ever more job-offers. Since they switch permanently to better-paying jobs, the wage they can get in their current job is the maximum of all wage offers that they have received

thus far. 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 rate δ_t. Then, the worker becomes unemployed and the selection process starts all over again. We refer to the time elapsing between the start of the labour market career and the first lay-off—the time elapsing between two consecutive lay-offs—as an employment cycle. A worker's current employment cycle has therefore started either at the last lay-off or —if a worker has never been laid off up till date— 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 Λ_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 λ_t is high) than during a bust (when λ_t is low). We define b as the termination date of the current job. Hence, Λ_a is the labour market time elapsed since the beginning of the employment cycle up until the date of job start, and Λ_b is the labour market time elapsed up until the date of job termination. Finally, let n denote the number of job offers received during the current employment cycle.

We assume that the job-offer arrival rate for employed job seekers λ_t varies proportionally to the job-offer arrival rate for unemployed job seekers λ_{ut}:

$$\lambda_t = \psi \lambda_{ut}. \tag{2}$$

The variation in the speed at which labour market time runs allows us to disentangle the regular return to labour market experience and the return to OJS.

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, where the rank is normalized such that it is uniformly distributed on $[0, 1]$. The proposition below specifies the relation between the elapsed labour market times Λ_a and Λ_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, Λ_a/Λ_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*

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

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

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

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 Λ_b measures the expected number of job offers from the start of the first job of the cycle until the moment of separation from the current job. Since λ_t is the job-offer arrival rate at time t, $\Lambda_b = \int_0^b \lambda_t dt$ is the accumulated arrival rate over that time interval. This process follows a Poisson distribution. Hence, the expected number of job offers is Λ_b. We should add one for the job offer that yielded the first job of the employement 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 $t = 0$.

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 the current job is the maximum value of w_{ia} of all job offers received up until time b. Job offers arrive proportionally to the value of λ_t. Suppose that n offers have arrived in the time interval $[0, b)$. They constitute n independent draws from the job-offer distribution, each associated with its own arrival time; a is the arrival time of the max. Each draw has an equal probability of being the maximum of these draws; a is thus uniformly distributed on the time scale of labour market time on the time interval $[0, b]$. Hence, Λ_a/Λ_b is distributed uniformly.

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 $\Lambda_b + 1$ draws from the uniform distribution is $1 - (\Lambda_b + 2)^{-1}$,[1] which explains the result. The relation is not exact, because we replace the actual number of job offers by its expectation, $\Lambda_b + 1$. Since the relation between the actual number of offers and the expected rank is non-linear, replacing the actual number by its expectation introduces a bias, which is of order $O\left(\Lambda_b^{-2}\right)$. Note that the difference between $(\Lambda_b + 2)^{-1}$ and Λ_b^{-1} is of order $O\left(\Lambda_b^{-2}\right)$, so the difference between the two can be ignored.

Next, consider the case of separation from the current job by means of a quit (see statement 4). The 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 expected number of offers is $\Lambda_b + 1$. In this case, the job that the worker holds up till time b is not the max, but is the second highest job offer, 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 is $1-2(\Lambda_b+2)^{-1}$.

[1] The expectation of the i-th order statistic for n draws of a uniform distribution is $i/(n+1)$.

Note that in both statements 3 and 4, $\mathrm{E}[F] \to 1$ if $\Lambda_b \to \infty$: if the selection of ever-better offers is allowed 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 δ_t, Δ_t. The intuition for this result is that the only thing that a lay-off achieves is that it stops the selection process and lets the worker start the process all over again. As long as the risk of a lay off has not materialized, it can be ignored. Statement 3 implies that the expected rank of the current job $\mathrm{E}[F]$ 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.

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 $\bar{w}_{it}$:

$$\bar{w}_{it} = \beta_i + \beta X_{it} + w_{ia} + \beta_x (t - a) + \epsilon_{it}.$$

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 $w_{ib} > w_{ia}$. Instead, the optimal strategy is to quit if $w_{ib} > w_{ia} + \beta_x (b - a)$. 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_{ia} currently held by the worker satisfies

$$w_{ia} = \arg\max_{s \in S} \left[w_{is} + \beta_x (a - s) \right].$$

The worker is prepared to move to a better job only if the gain in w_{ia} 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 β_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 Λ_a/Λ_b is skewed to the left relative to the uniform distribution, this shows that job-specific experience plays a

role. Hence, the distribution of Λ_a/Λ_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 Extreme Value Theory

Thus far, we have made no assumptions regarding the shape of the offer distribution except for the fact that $F(w)$ is differentiable. This section discusses how information of the conditional expectations can be used to identify the distribution. From Theorem 6.3.1 in Arnold, Balakrishnan, and Nagaraja (2008) we have the following:

Proposition 2 ***Nonparametric identification***

Let F and F' be arbitrary distributions. For $n = 1, 2, \cdots$ w_n and $w_{n'}$ denotes the maximum from n iid draws from F and F' respectively. Assume $E[w_n] = E[w_{n'}]$, $n = 1, 2, \cdots$ (assume finite). It follows that $F(w) = F(w)'$ for $w \in \mathbb{R}$.

The function $\mathrm{E}[w_n]$ can therefore be used to nonparametrically identify the distribution F. For sufficiently large values of b we can make strong predictions regarding the shape of $\mathrm{E}[w_{ia}|b]$ without knowing the exact distribution F. 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, showing that the normalised maximum of a large number of draws from a distribution converges to the Generalized Extreme Value (GEV) distribution for a large class of distributions. Proposition 3 states the relevant result from GEV distribution needed for our purpose.

Proposition 3 ***Generalized Extreme Value distribution***

1. *Let $x = \frac{w_n - \mu_n}{\sigma_n}$ be the maximum of n i.i.d. properly normalized draws w from some distribution $F(w)$, where μ_n and σ_n are normalizing constants. If the distribution of x converges to a stable distribution, then this distribution is the Generalized Extreme Value distribution $G(x)$*

which satisfies

$$G\,(x) = \exp[-t(x)] \;\; with \; t(x) = \begin{cases} (1+\xi x)^{-1/\xi} & if \; \xi \neq 0 \\ \exp\,(-x) & if \; \xi = 0 \end{cases}$$

2. *$G\,(x)$ has an unbounded upper support if $\xi \geq 0$ and an unbounded lower support if $\xi \leq 0$.*

3. *If F is unbounded above, $\xi \geq 0$.*

4. *If F is bounded above $\xi \leq 0$.*

5. *If F itself is the GEV distribution, then:*

   (a)

   $$E\,[w_n] = \begin{cases} \mu + \gamma + \sigma \ln n & if \; \xi = 0 \\ \mu_n & if \; \xi < 1 \vee \xi \neq 0 \\ \infty & if \; \xi \geq 1 \end{cases},$$

   where $\gamma = 0.577$ is Euler's constant and $\mu_n = \mu + \sigma \xi^{-1} \left[\left[\Gamma(1-\xi)\right] n^{\xi} - 1 \right]$;

   (b)

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

6. *If F is the normal distribution then $\xi = 0$ and*

$$\begin{aligned} \sigma_n &= \sigma(2\ln n)^{-1/2}, \\ \mu_n &= \sqrt{2\ln n - \ln(\ln n) - \ln(4\pi)}, \\ E\,[w_n] &\cong \mu_n + \gamma\sigma_n, \\ Var\,[w_n] &\cong \frac{\pi^2}{6}\sigma_n^2. \end{aligned}$$

Theorems 1.13 and 1.2.1 in De Haan and Ferreira (2007) prove that the only nondegenerate distribution that the normalized maximum can converge to is the GEV distribution. The book gives the conditions on the tail of the distribution F for which the distribution converges. Example 1.1.7 yields the normalising constants for the normal distribution.

The parameter ξ is referred to as the shape parameter. The cases $\xi < 0$, $\xi = 0$ and $\xi > 0$ correspond to the Weibull, Gumbel and Frechet distribution, respectively. Intuitively, the constants μ_n and σ_n ensure that the mean and variance, respectively, are well-behaved. 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 $\xi = 0$.

When $F(w)$ follows the Gumbel distribution, $F(w) = \exp\left(-e^{-(w-\mu)/\sigma}\right)$, 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. Then, the distribution of match quality for workers experiencing lay-offs is $F \exp[-\Lambda_b(1-F)]$ instead of F^{Λ_b+1}. 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 choosen 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.

The Normal distribution converges to the Gumbel distribution, but it does so slowly. For values of $\Lambda_b + 1 \leq 30$, however the expectation is approximated well by a second-order polynomial in $\ln(\Lambda_b + 1)$. Figure 2 plots the expected value of the maximum for the Normal distribution when the number of offers is Poisson distributed and $\beta_0 + \beta_1 \ln(\Lambda_b + 1) + \beta_2 \ln(\Lambda_b + 1)^2$. The approximation of the expected maximum for the normal distribution fits very well. This is important, as it suggests a simple test to use for the null hypothesis whereby the data is generated from the Normal distribution. Create a variable for the conditional expectation of the Normal distributon $\tilde{\Lambda}_b = \beta_0 + \beta_1 \ln(\Lambda_b + 1) + \beta_2 \ln(\Lambda_b + 1)^2$ and then include $\ln(\Lambda_b + 1)^2$, which should be zero under the null and equal to $-\beta_2$ for the case of the Gumbel distribution.

The speed of convergence to the GEV distribution differs between distributions; it is n^{-1} for the exponential distribution, whereas it is much lower for the normal distribution, only $(\ln n)^{-1}$. The speed of convergence depends on the shape of the right tail of the distribution. For example, after just

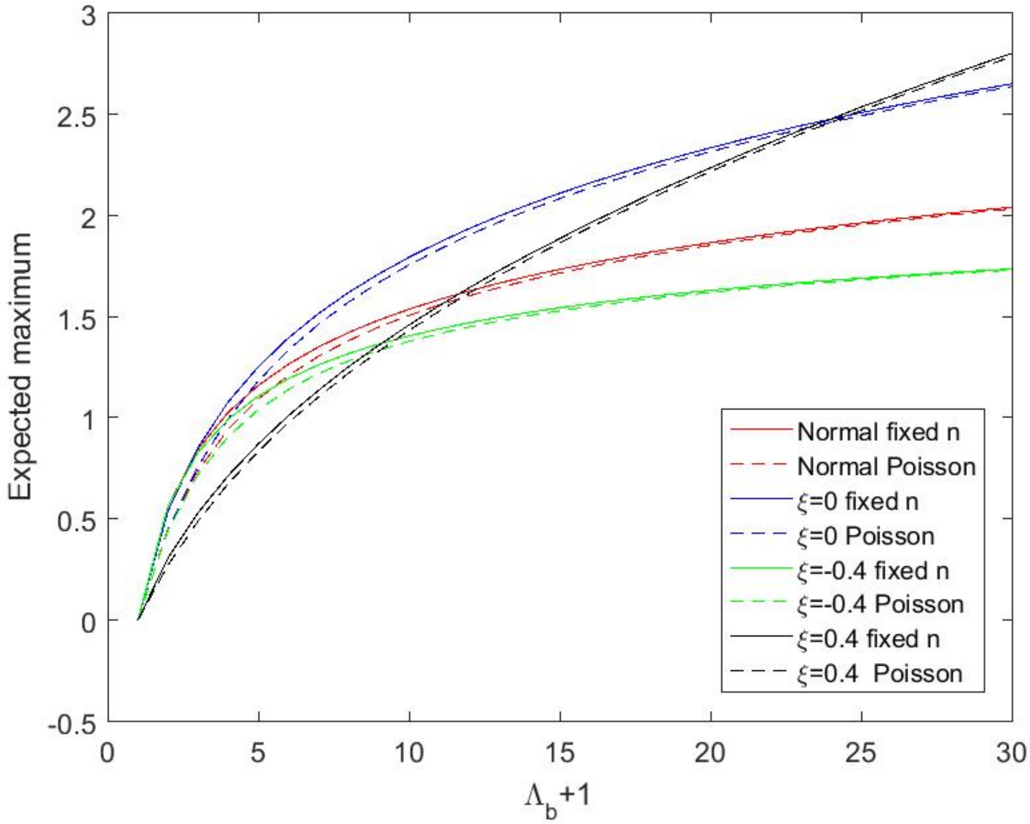

Figure 1: Expectation of the maximum as a function of the number of draws for four distribution

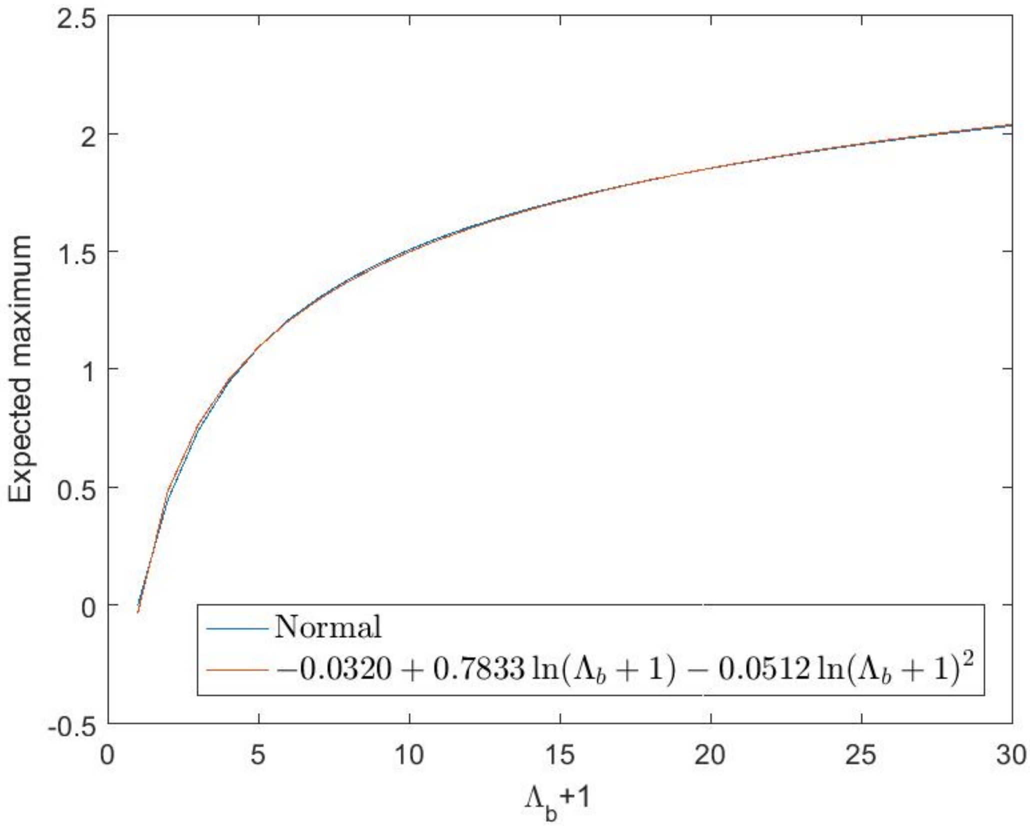

Figure 2: Approximation of the maximum of the Normal distribution

four job offers, $\Pr(F < 0.5) = 0.0625$ and $\mathrm{E}[F] = 0.80$. 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.

Jointly, Propositions 1 and 3 can be used for the derivation of expressions for $\mathrm{E}[w_{ia}|b,\text{lay-off}]$ and $\mathrm{E}[w_{ia}|b,\text{quit}]$. By Proposition 3, the expected value w_{ia} of the max of n offers satisfies

$$\mathrm{E}\left[w_{ia}|n\right] = \begin{cases} \sigma \ln n & \text{if } \xi = 0 \\ \gamma_\xi \sigma n^\xi & \text{if } \xi < 1 \vee \xi \neq 0 \end{cases},$$

When a job ends with a lay-off, $\mathrm{E}[n] = \Lambda_b + 1$ and value w_{ia} is the maximum of these offers; see Proposition 1[2]

$$\mathrm{E}\left[w_{ia}|\Lambda_b + 1,\text{lay off}\right] = \begin{cases} \sigma \ln\left(\Lambda_b + 1\right) + O\left(\Lambda_b^{-2}\right) & \text{if } \xi = 0 \\ \gamma_\xi \sigma \left(\Lambda_b + 1\right)^\xi + O\left(\Lambda_b^{\xi-2}\right) & \text{if } \xi < 1 \vee \xi \neq 0 \end{cases}. \quad (3)$$

When a job ends by a quit, we use the distribution of w_{ia} conditional on the fact that it is the second highest draw from n draws. The second highest draw is related to the highest draw by

$$\begin{aligned} \mathrm{E}[w_{ia}|n,\text{quit}] &= \int n(n-1) w f(w)(1-F(w))F(w)^{n-2} dw \\ &= n\mathrm{E}[w_{ia}|n-1] - (n-1)\mathrm{E}[w_{ia}|n]. \end{aligned}$$

For the case of the Gumbel distribution we have

$$\mathrm{E}[w_{ia}|n,\text{quit}] = \mathrm{E}[w_{ia}|n] + n \ln\left(\frac{n}{n-1}\right)\sigma = \mathrm{E}[w_{ia}|n] - \sigma + O(n^{-1}).$$

[2]The exact formula applies the expectation

$$\mathrm{E}\left[w_{ia}|b\right] = \sum_0^\infty \Pr\left(n|b\right) \mathrm{E}\left[w_{ia}|n\right] = \sum_0^\infty \left(n!\right)^{-1} \left(\Lambda_b + 1\right)^n e^{-\Lambda_b} \gamma_\xi \sigma n^\xi dn.$$

The formula in the text uses the first-order expansion $\mathrm{E}[w_{ia}|b] \cong \mathrm{E}[w_{ia}|\mathrm{E}\left[n\right]]$. A second-order expansion reads

$$\mathrm{E}[w_{ia|b}] \cong \gamma_\xi \sigma \left[\frac{(\Lambda_b + 1)^\xi}{\xi} + (\xi - 1)\frac{(\Lambda_b + 1)^{\xi-2}}{2}\Lambda_b\right].$$

Using the approximation for the number of offers from Proposition 1, statement 4 $\mathrm{E}[n] \cong \Lambda_b + 1$ yields an appealing result:

$$\mathrm{E}[w_{ia}|\Lambda_b,\text{quit}] \cong \sigma \ln(\Lambda_b + 1) - \sigma. \tag{4}$$

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

These relations can be used to estimate the value of ξ by Non-Linear-Least-Squares (NLLS). In practice, we estimate the equation by a grid search over different values of ξ, where we pick the value with the lowest residual variance. Note that none of the relations derived thus far depend on the lay-off rate δ_t. Hence, we can analyse the effect of OJS on wages without taking a stance on δ_t in general or its cyclical behavior in particular.

2.4 Applications and robustness

This subsection discusses a number of applications of this model. For the sake of convenience, we present expressions for the Gumbel distribution $\xi = 0$ only. These expressions tend to be simpler than those for other values of ξ. However, the expressions for the general case can be derived easily, using the formulas in Proposition 3.

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:

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

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 Λ_b, (ii) the number of job offers received, conditional on Λ_b, and (iii) the wage offer of the maximum, conditional on the number of offers. For the case of the Gumbel distribution, we obtain a particularly simple formula:

$$\begin{aligned} \mathrm{Var}\left[w_{ia}\right] &= \mathrm{Var}\left[E\left[w_{ia}|\Lambda_b\right]\right] + \mathrm{E}\left[\mathrm{Var}\left[\ln n|\Lambda_b\right]\right]\sigma^2 + \mathrm{Var}\left[w_{ia}|n\right] \qquad (5) \\ &\cong \left(\mathrm{Var}\left[\ln(\Lambda_b + 1)\right] + \mathrm{E}\left[\Lambda_b^{-1}\right] + \frac{\pi^2}{6}\right)\sigma^2, \end{aligned}$$

where $\pi = 3.14$. We apply the first-order approximation for the variance of the log Poisson distribution[3] using the fact that $\mathrm{Var}[w_{ia}|n] = \frac{\pi^2}{6}\sigma^2$ does not depend on n for the Gumbel distribution (see statement 2 of Proposition 3). Estimates for the first two terms can be obtained from the data. The value of σ is derived from the estimation of equation (3).

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 last job of the current employment cycle (for which $\mathrm{E}[n] = \Lambda_b + 1$) with the expected log wage in the first job of the new employment cycle (for which $n = 1$). The expected loss in log wages can be calculated as

$$\mathrm{E}[w_{ia}|n] - \mathrm{E}[w_{ia}|n=1] \cong \mathrm{E}[w_{ia}|\mathrm{E}[n]] - \mathrm{E}[w_{ia}|n=1]$$
$$= \int_0^\infty \Pr(t)\,\sigma \ln(\Lambda_t + 1)\,dt = -\sigma \exp(\delta/\lambda)\,\mathrm{Ei}(-\delta/\lambda) \tag{6}$$

since $\Pr(t) = \delta \exp[-\Delta_t]$.

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 seeker, denoted X_i.[4] We assume that the arrival rate follows the well-known proportional hazard model: $\lambda_t \exp\left(\theta' \widetilde{X}_i\right)$, $\widetilde{X}_i$ is the deviation of X_i from its mean; $\exp\left(\theta' \widetilde{X}_i\right)$ is the baseline hazard rate, and θ is a parameter vector. Compared to equation (3), we should replace Λ_b by $\Lambda_b \exp\left(\theta' \widetilde{X}_i\right)$. Again,

---

[3]

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

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

[4] In the current specification, X_i can include the fixed worker effect β_i; it cannot include experience, since then λ_t would depend on time not only due to the business cycle but also due to the return to experience.

focussing on the Gumbel distribution for the sake of convenience, we obtain

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

Differences in the job-offer arrival rate between individuals are therefore absorbed in the fixed effects up to a term of order $\mathrm{Var}[X_i]$. As long as the coefficient of variation of X_i is small relative to Λ_b, variation in the hazard rate is absorbed in the term β_i in equation (1).

2.5 The identification of λ_{ut} and ψ

The job-offer arrival rate λ_t is a critical input for the exercise discussed above. Although, we observe accepted job offers by counting job-to-job transitions, we have no reliable measure of rejected offers. This problem does not plague the job-offer arrival rate for unemployed job seekers λ_{ut}, since the unemployed accept all job offers. Hence, we derive an estimate of λ_t from the observed transition rate from unemployment to employment and then apply equation (2). This subsection describes how we can estimate λ_{ut} and ψ. The transition rate λ_{ut} can be estimated from data on the transitions from unemployment to employment. These data are readily available in most panel data. We minimize the impact of the selection bias introduced by unobserved heterogeneity in job-finding rates by restricting the analysis to these job seekers who have been unemployed for less than five weeks. The calculation of the transition rates is described in detail in the empirical section. We present a new method to back out ψ from data on job-to-job transitions, using the expected duration of the first job and the average length of the employment cycle b for the subsequent jobs. For this derivation, we rely on a steady-state argument, where labour market time runs at a constant rate. Hence, we drop the suffix t of λ_t and δ_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 $\delta + \lambda\overline{F}$ where $\overline{F} \equiv 1 - F$ is the complement of F Hence, the expected duration of a job conditional on its rank is $(\delta + \lambda\overline{F})^{-1}$. Since the rank of the first job is a random draw from the uniform distribution, its

expected duration satisfies

$$\mathrm{E}\left[b|1^{st}\text{ job in emp.cycle}\right] = \int_0^1 (\delta+\lambda)^{-1} dF = \lambda^{-1}\ln\left(1+\delta/\lambda\right). \quad (7)$$

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 denstity is comprised of three parts: (i) the fraction $F\exp\left[-\left(\delta+\lambda\overline{F}\right)a\right]$ of workers remaining at a with rank less than F, (ii) the arrival rate λ of an offer at a, and (iii) the probability $\left(\delta+\lambda\overline{F}\right)\exp\left[-\left(\delta+\lambda\overline{F}\right)(b-a)\right]$ that a match ends at b conditional on it having started at a. Hence, this density is proportional to

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

In the second line, a drops out. We can ignore the job-offer arrival rate λ, 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:

$$\Pr(F,b)\propto F\exp\left[-\left(\delta+\lambda\overline{F}\right)b\right]\left(\delta+\lambda\overline{F}\right)b,$$

Hence:

$$\mathrm{E}\left[b|\text{subseq.jobs}\right] = \int_0^1\int_0^\infty b\Pr(F,b)\;db\;dF = \frac{2}{\lambda}\frac{\lambda/\delta-\ln\left(1+\lambda/\delta\right)}{\left(1+\delta/\lambda\right)\ln\left(1+\lambda/\delta\right)-1}. \quad (8)$$

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

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[5]. We exclude job observations in cases when hours worked per week are less than 15 and when job spells lasted shorter than four weeks or started before 1979[6]. When there are multiple jobs, the primary job is defined as the job with highest number of hours[7]. Jobs with inconsistencies in their start and end date are adjusted or removed[8]. 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[9].

For the construction of the variable Λ_b, we have to categorize job terminations into either quits (belonging to the same employment cycle) or lay-offs (starting a new cycle). We follow Barlevy (2008) and Hagedorn and Manovskii (2013), 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).[10] 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.

[5]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.

[6]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.

[7]If hours are the same, we used the average hours for that job spell and the length of the job spell to determine which is the main job.

[8]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.

[9]We use the "adjusted"schooling variable.

[10]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.

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 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[11]. 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.[12] 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 apply the method discussed in Section 2.5 to estimate the efficiency of OJS relative to search as an unemployed ψ. From our data we calculate two statistics, where the unit of observation is a job:

$$\begin{aligned} \mathrm{E}[\Lambda_b|1^{st} \text{ job in emp.cycle}] &= 1.6, \\ \mathrm{E}[\Lambda_b|\text{subsequent jobs}] &= 4.6. \end{aligned}$$

Using these statistics and the mean value of λ_u, we obtain the following values for the monthly transition rates and for the relative efficiency of OJS:

$$\begin{aligned} \lambda_u &= 0.4, \\ \lambda &= 0.08, \\ \delta &= 0.02, \\ \psi &= 0.2. \end{aligned}$$

[11] We match the monthly CPS using variables suggested by Drew and Warren (2014). In addition, we use race and age as extra controls.

[12] 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	0.800	0.793	0.793	0.789
	(0.418)	(0.421)	(0.421)	(0.426)
Fraction high	0.316	0.395	0.395	0.359
	(0.465)	(0.489)	(0.489)	(0.480)
$\bar{w}_{it}$	-2.841	-2.660	-2.660	-2.681
	(0.538)	(0.554)	(0.554)	(0.558)
$\ln(\Lambda_b + 1)$	0.709	1.300	1.300	1.345
	(0.598)	(0.690)	(0.690)	(0.879)
$\ln(\Lambda_{Tb} + 1)$	1.569	1.952	1.952	1.968
	(0.837)	(0.666)	(0.666)	(0.803)
$\ln(\lambda T_b + 1)$	0.716	1.307	1.307	1.352
	(0.596)	(0.682)	(0.682)	(0.874)
Individuals	2572	1470	607	2582
Jobs	12624	2254	991	15869
Observations				33387

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.

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 this job. This implies that Λ_a/Λ_b for all jobs other than the first job of an employment cycle should be distributed uniformly. Figure 3 presents the kernel estimate of the density of Λ_a/Λ_b, separately for low- and high-skilled 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, Λ_a/Λ_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 4. 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.

The results of this test provide a strong and quite unexpected confirmation of the simple model 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_{ia}. Our results show the opposite.

3.3 The shape of the offer distribution

The uniform distribution of Λ_a/Λ_b justifies the application of Proposition 3 for the analysis of the evolution of the match-specific component w_{ia} over the employment cycle. This section presents the tests of the model as well as an estimate for the shape of the offer distribution. All of our regressions include tenure and experience up to a third order and quadratics in years of education and time, as well as dummies for region, marriage, and urban versus rural location of the job. 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 the discussion related to equation 4). The inclusion of a third order polynomial in tenure is debatable. Strictly speaking, 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

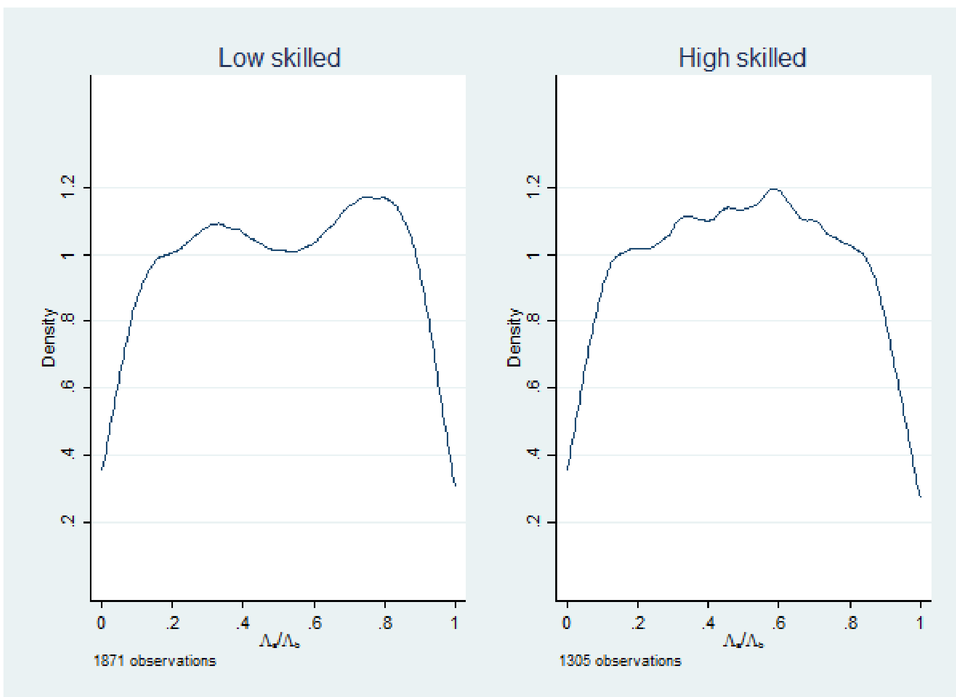

Figure 3: Test of the arrival rate of the maximum

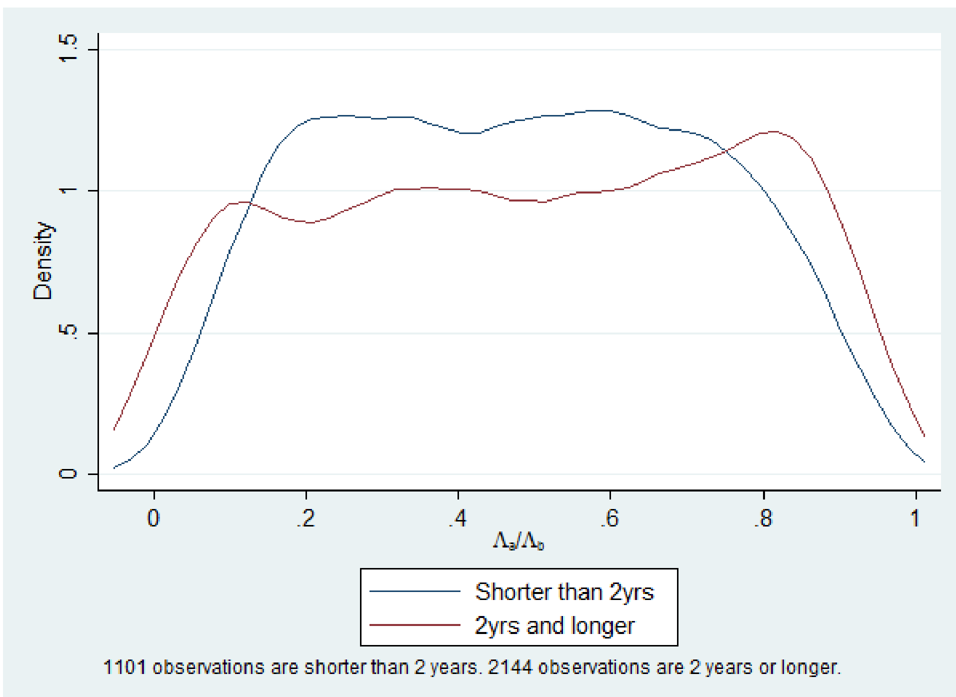

Figure 4: Test of the arrival rate of the maximum by employment length

of the job. If there would be a tenure profile in wages, then the optimal strategy of workers would no longer involve accepting each job offer with a higher value of w_{ia} than the initial value of w_{ia} in the current job. The results regarding the distribution of Λ_a/Λ_b presented in the previous section suggest that job-specific returns are unimportant. We nevertheless include this third-order polynomial to prevent that the term for the expected number of job offers Λ_b captures the effect of the omitted variable 'tenure'. We shall discuss the effect of tenure later on, in Section 3.5.

We estimate the non-linear equation

$$\bar{w}_{it} = \beta_i + \beta' X_{it} + \sigma \frac{\Lambda_b^{\xi} - 1}{\xi} + \varepsilon_{it}$$

for different samples. There are two reasons to be cautious about this method. First, we use the expected rather than the actual number of job offers n in our derivations. Since the relation between n and w_{ia} is non-linear, this introduces a bias, particularly for low values of n. Second, the NLLS procedure for the estimation of ξ is not unbiased itself. To quantify these drawbacks we therefore generate wage data drawn from the GEV distribution for different values $\xi \in \{-0.4, 0, 0.4\}$ and from the normal distribution and run our regression on these simulated data. We generate data by applying the following procedure. We use the joint distribution of Λ_b and the type of job separation (lay-off or quit) from the data. We draw match quality from the conditional distribution of F given Λ_b and the type of separation; the formula is presented in the derivation in Appendix A. For each distribution and simulation the scale parameter is chosen such that the coefficient from a regression on $\ln(\Lambda_b + 1)$ is the same as in the data. We draw a normally distributed error term ε_{it} for each observation with a scale parameter such that the variance of the sum of the generated data matches the residual variance from a regression on just the controls. The parameter ξ is then estimated by a grid search as described. We then repeat this procedure a one hundred times, each time obtaining an estimate of ξ. The results appear in Table 2. Column 1 presents the results from the grid search for different values of ξ and columns 2, 3, 4 and 5 show the estimation on the simulated data, with the standard deviation of the estimates appear in parentheses. Examining the simulated data reveals that our estimation procedure has a small upward bias. The estimation on a sample containing longer jobs suggests that the distribution has a fat right tail. The estimated coefficients when we restrict the data for jobs with T_b greater than 1/2, 1 or 2 years are all very close to

Table 2: Estimation of the shape parameter

	Data	$\xi = -0.4$	$\xi = 0$	$\xi = 0.4$	Normal
$\xi_{\text{All jobs}}$	-0.001	-0.179	0.123	0.490	-0.072
	(.)	(0.073)	(0.077)	(0.139)	(0.074)
$\xi_{T_b \geq 1/2 \text{ year}}$	0.080	-0.254	0.092	0.485	-0.127
	(.)	(0.107)	(0.113)	(0.186)	(0.109)
$\xi_{T_b \geq 1 \text{ year}}$	0.137	-0.352	0.048	0.473	-0.200
	(.)	(0.131)	(0.128)	(0.185)	(0.128)
$\xi_{T_b \geq 2 \text{ years}}$	0.062	-0.422	0.019	0.426	-0.241
	(.)	(0.166)	(0.192)	(0.298)	(0.175)
$\xi_{T_b \leq 10 \text{ years}}$	-0.266	-0.051	0.179	0.440	0.022
	(.)	(0.118)	(0.123)	(0.191)	(0.119)

Notes: Mean coefficient and standard deviations of estimates in parentheses. For the different samples of the data the ξ parameter is estimated using a grid search over different values and generated data using the four distributions and the empirical distribution of Λ_b.

the mean estimate from the Gumbel distribution. We reject the hypothesis that the shape parameter ξ is -0.4 or $\xi = 0.4$.

When we estimated the shape parameter we made a parametric assumption on the function $E[w|\Lambda_b]$. We now test this parametric assumption by including higher-order terms of $\ln(\Lambda_b + 1)$. The predicted values for the different regressions are presented in the left panel of Figure 5 for the central 90 percentiles of the distribution of Λ_b. The results show that the conditional expectation is approximated well by the first order term. The right panel of Figure 5 presents the results from a nonparametric regression on the residuals from the regression with only the first order term. The results confirm that the regression approximated the conditional expectation well. Hence, we cannot reject the hypothesis that the offer distribution is Gumbel.

We perform a separate test for the Normal distribution. From Figure 2 we take the coefficients from the polynomial in $\ln(\Lambda_b + 1)$ to create the the conditional expectation of the Normal distribution which we call $E[w|\text{Normal}, \Lambda_b]$. We then run a regression using $E[w|\text{Normal}, \Lambda_b]$ and $\ln(\Lambda_b + 1)^2$ for all data

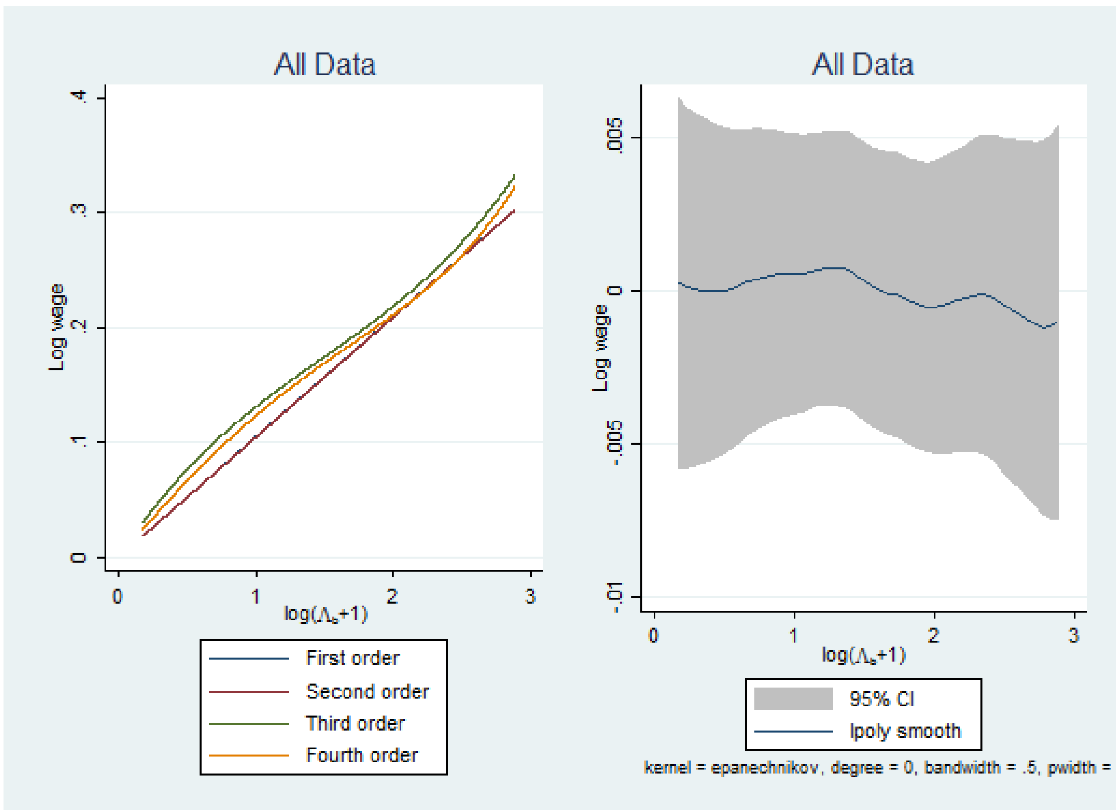

Figure 5: Robustness test for shape parameter

Table 3: Test for the Normal distribution

	All Data	$T_b \geq 1/2$ year	$T_b \geq 1$ year
u_t	-0.014***	-0.013***	-0.011***
	(0.001)	(0.002)	(0.002)
$E[w \mid \text{Normal}, \Lambda_b]$	0.133***	0.113***	0.100***
	(0.016)	(0.023)	(0.032)
$\ln(\Lambda_b + 1)^2$	0.007**	0.009**	0.012**
	(0.003)	(0.004)	(0.005)
Quit	-0.041***	-0.040***	-0.036***
	(0.006)	(0.006)	(0.007)
Observations	33387	27788	23780
R^2	0.645	0.669	0.697

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 the Normal and Gumbel distribution.

and for values of T_b greater than 1/2 and 1 year. Under the null hypothesis of a Normal distribution the coefficient of $\ln(\Lambda_b + 1)^2$ should be zero —whereas under the Gumbel distribution it should be positive. The results appear in table 3. The null-hypothesis of the offer distribution being normal is clearly rejected. Moreover, the ratio of the coefficient for second order term relative to the coefficient on $E[w|\text{Normal}, \Lambda_b]$ fits the ratio of the theoretical coefficients in Figure 2 very well. Hence, the second term offsets the second effect included in the term $E[w|\text{Normal}, \Lambda_b]$, so that the estimated shape is again consistent with the Gumbel distribution.

Apart from the convenience of the Gumbel distribution, this conclusion has an important implication: that the offer distribution has an unbounded upper support. This 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)), it cannot be the full story. In all subsequent regressions, we use the Gumbel distribution and hence logarithmic transformation of $\Lambda_b + 1$.

3.4 Further test 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 high-skilled workers. Across all regressions, the effect of OJS tends to be one and half times bigger for high- rather than low-skilled 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. As discussed before, the dummy for quits should be negative and the magnitude should be the same as the standard deviation (see equation (4)). In all estimations the dummy has the right sign, but it tends to be smaller than the model predicts.

Our first test checks whether the coefficient σ for $\ln(\Lambda_b + 1)$ is stable across subsequent employment cycles, as is predicted by the model. In Table 11, 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 σ is stable across employment cycles. Though the coefficient for $\ln(\Lambda_b+1)$ is somewhat lower for later cycles, the order of magnitude is the same and the coefficient is highly significant for each cycle. This is a strong confirmation of our model. In Appendix B we present separate estimates for each of the first eight cycles. These results yield the same conclusion.

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

Table 4: Estimation for the different employment cycles

	All Data	Low-Skilled	High-Skilled
u_t	-0.014***	-0.015***	-0.012***
	(0.001)	(0.002)	(0.003)
$\ln(\Lambda_b + 1)1(\text{cycle} \leq 3)$	0.118***	0.090***	0.146***
	(0.005)	(0.007)	(0.009)
$\ln(\Lambda_b + 1)1(3 < \text{cycle} \leq 6)$	0.107***	0.091***	0.128***
	(0.005)	(0.007)	(0.010)
$\ln(\Lambda_b + 1)1(6 < \text{cycle})$	0.077***	0.077***	0.103***
	(0.006)	(0.008)	(0.013)
Quit	-0.038***	-0.053***	-0.012
	(0.006)	(0.007)	(0.011)
Observations	33387	20886	11722
R^2	0.646	0.565	0.676

Standard errors in parentheses

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

Table 5: Robustness test for OJS

	All Data	Low-Skilled	High-Skilled
u_t	-0.015***	-0.016***	-0.012***
	(0.001)	(0.002)	(0.003)
$\ln(\Lambda_b+1)$	0.092***	0.063***	0.144***
	(0.007)	(0.009)	(0.013)
$\ln(\Lambda_{Tb}+1)$	0.028***	0.054***	-0.019
	(0.011)	(0.013)	(0.019)
Quit	-0.042***	-0.057***	-0.014
	(0.006)	(0.007)	(0.011)
Observations	33387	20886	11722
R^2	0.646	0.565	0.675

Standard errors in parentheses

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

this, we run a regression including both variables: see Table 5. For both low- and high-skilled workers, the number of offers during the current cycle turns out to be the most important and highly significant, although for the low-skilled the number of offers during the whole career matters, as well. This supports the OJS model.

The effect of the number of offers during the whole career 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. Table 6 runs the same regressions as in Table 5 on simulated data, where we assume that a proportion of the separations classified as lay-offs are in fact quits and therefore belong to the same employment cycle. Any layoff we assume with probability $p \in \{10\%, 20\%, 30\%\}$ to have been a quit. The number of job offers are taken to be deterministic over the job spells. For the complete spell the number of offers received is then $\Lambda_b + 1$ and during job k the number of offers received is $\Lambda_b^k - \Lambda_b^{k-1}$. We then draw the match quality for the last job and then for the previous jobs, recursively. The results suggest that a moderate amount of misclassification can explain the estimated coefficients in Table 5. The title in Table 6 refers to the fraction of the job changers that are misclassified as lay-offs.

In Table 7 we include —next to $\ln(\Lambda_b+1)$— the variable $\ln(\lambda T_b + 1)$,

Table 6: Simulation of misclassification

	10%	20%	30%
$\log(\Lambda_b + 1)$	0.104	0.101	0.097
	(0.004)	(0.004)	(0.004)
$\log(\Lambda_T + 1)$	0.002	0.009	0.019
	(0.010)	(0.010)	(0.010)

Notes: Mean coefficients and standard deviations from 1000 simulations in parentheses.

where λ measures the average number of offers per unit of calendar time and where T_b measures calendar time in the current employment cycle. Hence, Λ_b and λT_b measure the same time episode, on average in the same units, but on different clocks; Λ_b takes account of the business cycle fluctuations in the job-offer arrival rate λ_t. By entering both variables 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, $\ln(\lambda T_b + 1)$, 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.

The results provide support for the OJS model. Columns 1-3 report the results with both calendar- and labour market time. Labour market time clearly outperforms calendar time. The coefficients on calendar time have the wrong sign, though it is only significant when combining low- and high-skilled workers. Columns 4-6 report the results with only calendar time. This yields similar estimates for σ as when one only includes labour market time; the coefficient on $\ln(\lambda T_b + 1)$ 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 σ of similar magnitude.

In Table 8 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 high-skilled workers. This is consistent with models 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 equation (11) and (11)), the expected

Table 7: Robustness test for OJS

	Calendar and labour market time			Only calendar time		
	All Data	Low-Skilled	High-Skilled	All Data	Low-Skilled	High-Skilled
u_t	-0.012***	-0.013***	-0.011***	-0.016***	-0.016***	-0.014***
	(0.002)	(0.002)	(0.003)	(0.001)	(0.002)	(0.003)
$\ln(\Lambda_b + 1)$	0.259***	0.194**	0.274*			
	(0.078)	(0.096)	(0.146)			
$\ln(\lambda T_b + 1)$	-0.155**	-0.108	-0.140	0.104***	0.086***	0.135***
	(0.078)	(0.096)	(0.146)	(0.005)	(0.006)	(0.008)
Quit	-0.041***	-0.054***	-0.015	-0.041***	-0.054***	-0.015
	(0.006)	(0.007)	(0.010)	(0.006)	(0.007)	(0.010)
Observations	33387	20886	11722	33387	20886	11722
R^2	0.645	0.565	0.675	0.645	0.565	0.675

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Estimation for subgroups

	Low-Skilled		High-Skilled	
	Rural	Urban	Rural	Urban
u_t	-0.009**	-0.017***	0.000	-0.013***
	(0.004)	(0.002)	(0.008)	(0.004)
$\ln(\Lambda_b + 1)$	0.074***	0.073***	0.020	0.138***
	(0.015)	(0.008)	(0.033)	(0.012)
Quit	-0.098***	-0.050***	0.026	-0.025*
	(0.017)	(0.008)	(0.054)	(0.013)
Observations	3786	13589	1124	8036
R^2	0.584	0.572	0.705	0.667

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Regressions using the reason for separation

	All Data	Low-Skilled	High-Skilled
u_t	-0.014***	-0.015***	-0.010***
	(0.002)	(0.002)	(0.004)
$\ln(\Lambda_b+1)1(Fired)$	0.081***	0.075***	0.100***
	(0.007)	(0.009)	(0.016)
$\ln(\Lambda_b+1)1(Quit)$	0.102***	0.079***	0.120***
	(0.008)	(0.010)	(0.013)
Quit	-0.115***	-0.097***	-0.105***
	(0.013)	(0.015)	(0.030)
Observations	18991	11841	6730
R^2	0.707	0.608	0.747

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

productivity of a match depends on whether it ends with a quit or a lay-off. The difference in intercept should be equal to σ. As a further test we run the regressions of wages on labour market history for both groups separately. Table 9 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 increase in wages with Λ_b is small. The difference in means captured by the dummy for quits is close to the estimate of σ.

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 $(\Lambda_b + 1)^{-1}$ (see the discussion related to equation (5)). 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 an heavily right tailed distribution, we exclude the largest 5% of the squared residuals to reduce the impact of outliers. The results appear in Table 10. 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 high in the

Table 10: Regressions for the variance of wages

	All Data	Low-Skilled	High-Skilled
$(\Lambda_b + 1)^{-1}$	0.026***	0.018***	0.041***
	(0.003)	(0.003)	(0.005)
Observations	31718	19842	11136
R^2	0.026	0.021	0.042

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

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, as then the variance of wages conditional on the number of offers n is decreasing in n (see Proposition 3).

3.5 Applications of the OJS model

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

$\mathrm{Var}[\mathrm{E}[\ln n]]$	$\mathrm{Var}[\ln(\Lambda_b + 1)]\,\sigma^2$	0.0077
$\mathrm{Var}[\ln n \mid \mathrm{E}[\ln n]]$	$\mathrm{E}\left[\Lambda_b^{-1}\right]\sigma^2$	0.0016
$\mathrm{Var}[w_{ia} \mid n]$	$\frac{\pi^2}{6}\sigma^2$	0.0181
$\mathrm{Var}[w_{ia}]$	$\left(\mathrm{Var}[\ln(\Lambda_b + 1)] + \mathrm{E}\left[\Lambda_b^{-1}\right] + \frac{\pi^2}{6}\right)\sigma^2$	0.0275
$\mathrm{Var}[\overline{w}_{it}]$		0.2970
Share	$\mathrm{Var}[w_{ia}]\,/\mathrm{Var}[\overline{w}_{it}]$	9.2%

The result from Table 9 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 the simple expression (equation 6) for the average wage loss in terms of just two parameters, λ/δ and σ. λ/δ 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 λ/δ is, the

better are the outstanding matches compared to the average match a worker coming out of unemployment gets. σ measures the scale of the distribution. For a given drop in match quality, the loss in wages increases in σ. Using $\sigma = 8\%$ and our estimate of $\lambda/\delta \cong 4$, we see that a loss of match quality can explain an average wage loss after firm lay-off 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)), suggesting that the loss of firm specific human capital or other factors must also play a role. 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 last twice as long in recessions.

Workers experience an increase in expected wages over the life cycle. There are several potential explanations for this phenomenon. 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 can 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. If we regress wages on tenure without controlling for match quality, the tenure profile will be upward biased (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 gives 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

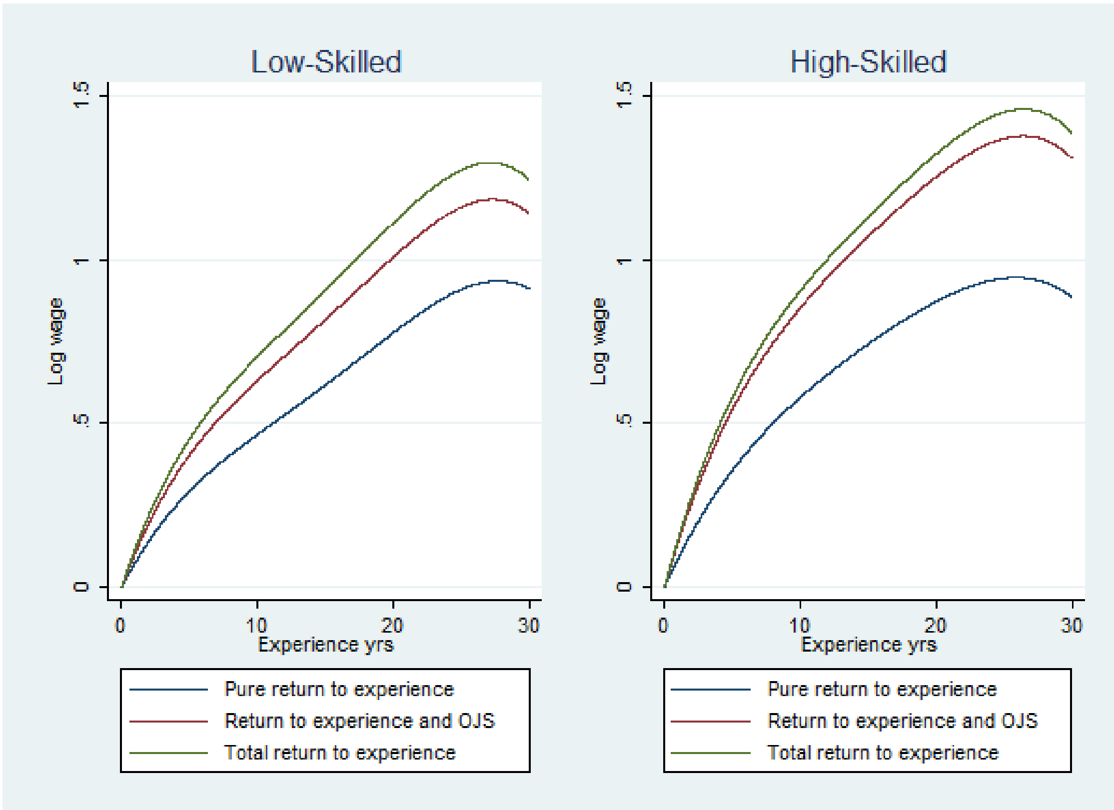

Figure 6: Experience Profile with and without controlling for OJS

by regressing wages on fourth-order polynomials in experience and tenure including $\ln(\Lambda_b+1)$. The estimated coefficients on the polynomial yeilds the pure experience profile. These estimates appear in Figure 6. 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.

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 7. The results suggest that most of the raw tenure profile is due to survival bias. This result explains why we find that Λ_a/Λ_b is uniformly distributed. The return

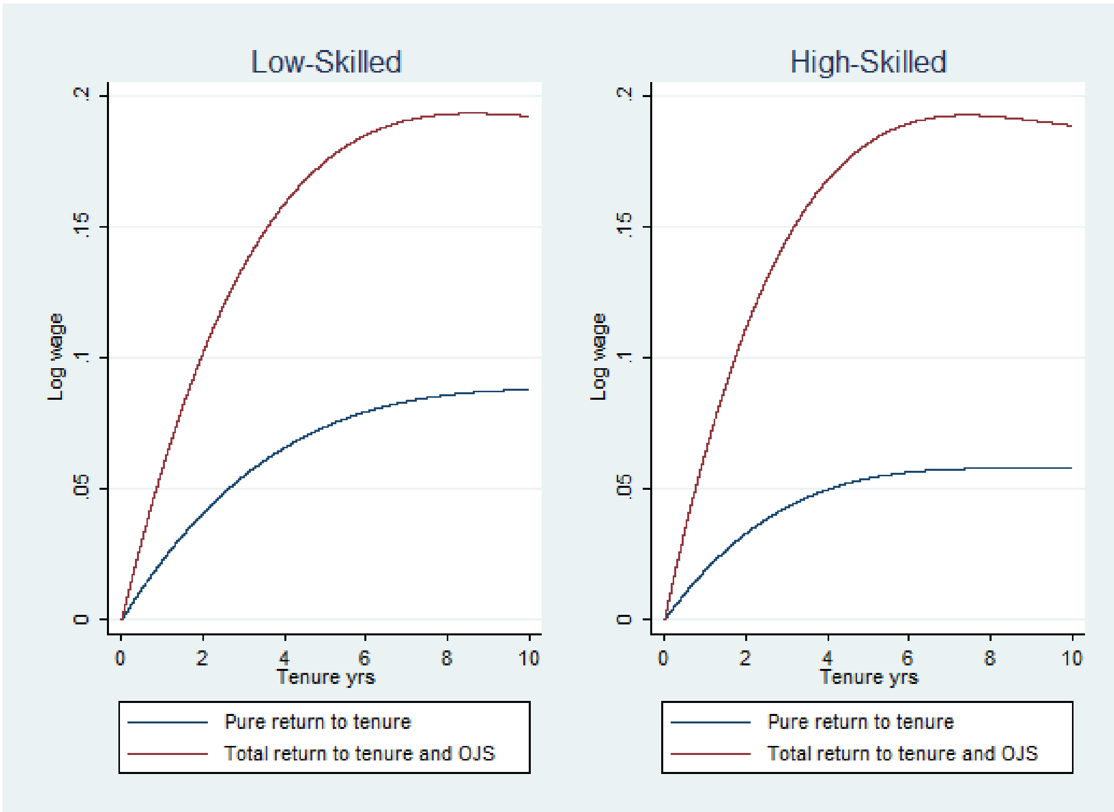

Figure 7: Tenure Profile with and without controlling for OJS

to job-specific experience and the true tenure profile in wages are apparently not that important.

References

Altonji, J. G., A. A. Smith, and I. Vidangos (2013): "Modeling earnings dynamics," *Econometrica*, 81, 1395–1454.

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.

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.

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.

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.

Nagypal, E. (2006): "On the extent of job-to-job transitions," 2006 Meeting Papers 10, Society for Economic Dynamics.

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.

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 $[0, b]$:

   First, a worker receives one offer at the start of an employment cycle. Next, the accumulated arrival rate of job offers during period $(0, b)$ is Poisson distributed with parameter Λ_b. The expectation of the Poisson distribution is equal to the arrival rate.

   For the case of a job ending by lay-off the expected number of job offers is therefore $\Lambda_b + 1$ on the interval $[0, b]$.

   For the case of a quit, the problem is slightly more involved, since the likelihood of an acceptable offer coming in a time b depends on the number of offers received previously. The probability of $n-2$ offers during the interval $(0, b)$ is $\exp(-\Lambda_b) \frac{\Lambda_b^{n-2}}{(n-2)!}$. The probability of an acceptable offer arriving at time b conditional on $n-1$ previous offers is $n^{-1}\lambda_b$; λ_b is the arrival rate of job offers and n^{-1} is the probability that the offer is acceptable. Because the worker will have received n offers (one at time 0, $n-2$ in the interval $(0, b)$, and one at time b), the probability is that the last offer is the highest is n^{-1}. The joint probability is therefore $n^{-1}\lambda_b \exp(-\Lambda_b) \frac{\Lambda_b^{n-2}}{(n-2)!}$. Taking expectations over n and using the normalizing constant yields

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

2. Λ_a/Λ_b is uniformly distributed on the unit interval $[0, 1]$:

   The distribution of outstanding match quality $\underline{F}$ at time t for employed workers satisfies

$$\Pr\left(\underline{F} < F | t\right) = F \exp\left[-\Lambda_t \bar{F}\right], \tag{9}$$

   where $\bar{F} \equiv 1 - F$. Let a be the time at which a job started and let b be the time at which this job will end. Conditional on a job starting at a with match quality F, the distribution of end dates b satisfies

$$\Pr(b | F, a) = \left[\delta_b + \lambda_b \bar{F}\right] \exp\left[-\left(\Lambda_b - \Lambda_a\right) \bar{F} - \left(\Delta_b - \Delta_a\right)\right]. \tag{10}$$

Now we want to derive the joint probability of starting at a and ending at b. Using equation (9) for $t = a$ to derive the distribution of the quality in the previous job, we can therefore write the joint probability of the job quality F in the current job, its termination time b and its starting time a as follows:

$$\begin{aligned}\Pr(F, b, a) &= C_1 \Pr\left(\underline{F} < F | a\right) \Pr\left(F | \text{offer at } a\right) \Pr\left(\text{offer at } a\right) \Pr(b | F, a) \\ &= C_2 \left[\delta_b + \lambda_b \bar{F}\right] \lambda_a \exp\left[-\Lambda_b \bar{F} - \Delta_b\right] F.\end{aligned}$$

Note that a only enters the expression via λ_a. Integrating over F gives $G(b, a) = \lambda_a C(b)$. In order to get the density of Λ_a we use $dt = \lambda_t d\Lambda_t$ which gives

$$\Pr(b, a) = C_2(b).$$

Therefore the distribution of a is uniform.

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

   The distribution function of F after one initial offer at time 0 and $\mathrm{E}[n] = \Lambda_b$ offers in $(0, b)$ is

$$\Pr\left(\underline{F} < F\right) = F \exp\left(-\Lambda_b \bar{F}\right).$$

   Hence, the density function reads

$$\Pr\left(\underline{F} = F\right) = \left(\Lambda_b F + 1\right) \exp\left(-\Lambda_b \bar{F}\right)$$

   Using this, the expectation of F can be written as

$$\begin{aligned}\mathrm{E}(F | b, \text{lay-off}) &= \int_0^1 F\left(\Lambda_b F + 1\right) \exp\left(-\Lambda_b \bar{F}\right) dF \\ &= 1 - \Lambda_b^{-1} + \Lambda_b^{-2}\left[1 - \exp\left(-\Lambda_b\right)\right] = 1 - \Lambda_b^{-1} + O\left(\Lambda_b^{-2}\right).\end{aligned}$$

4. Expected rank for jobs ending in a quit:

$$\begin{aligned}\mathrm{E}(F | b, \text{quit}) &= \int_0^1 \frac{\Lambda_b \bar{F}}{1 - \Lambda_b^{-1}\left[1 - \exp\left(-\Lambda_b\right)\right]} \left(\Lambda_b F + 1\right) \exp\left(-\Lambda_b \bar{F}\right) dF \\ &= 1 - 2 \frac{\Lambda_b + \left(\Lambda_b + 2\right) \exp\left(-\Lambda_b\right) - 2}{\Lambda_b + \exp\left(-\Lambda_b\right) - 1} \Lambda_b^{-1} = 1 - 2\Lambda_b^{-1} + O\left(\Lambda_b^{-2}\right).\end{aligned}$$

B Estimation results by employment cycle

Table 11: Estimation for the different employment runs

	All Data	Low-Skilled	High-Skilled
u_t	-0.014***	-0.015***	-0.012***
	(0.001)	(0.002)	(0.003)
$\ln(\Lambda_b + 1)1(\text{cycle } 1)$	0.116***	0.090***	0.147***
	(0.008)	(0.011)	(0.014)
$\ln(\Lambda_b + 1)1(\text{cycle } 2)$	0.118***	0.088***	0.148***
	(0.006)	(0.008)	(0.010)
$\ln(\Lambda_b + 1)1(\text{cycle } 3)$	0.122***	0.092***	0.152***
	(0.006)	(0.008)	(0.011)
$\ln(\Lambda_b + 1)1(\text{cycle } 4)$	0.111***	0.085***	0.147***
	(0.006)	(0.008)	(0.012)
$\ln(\Lambda_b + 1)1(\text{cycle } 5)$	0.100***	0.092***	0.111***
	(0.007)	(0.008)	(0.014)
$\ln(\Lambda_b + 1)1(\text{cycle } 6)$	0.116***	0.107***	0.114***
	(0.008)	(0.011)	(0.015)
$\ln(\Lambda_b + 1)1(\text{cycle } 7)$	0.098***	0.095***	0.133***
	(0.009)	(0.011)	(0.017)
$\ln(\Lambda_b + 1)1(\text{cycle } 8)$	0.080***	0.076***	0.085***
	(0.009)	(0.011)	(0.018)
$\ln(\Lambda_b + 1)1(8 < \text{cycle})$	0.063***	0.071***	0.070***
	(0.008)	(0.009)	(0.017)
Quit	-0.038***	-0.054***	-0.012
	(0.006)	(0.007)	(0.011)
Observations	33387	20886	11722
R^2	0.646	0.565	0.676

Standard errors in parentheses

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