Search Capital and the Wages of Young Men*

Gadi Barlevy
Department of Economics
Northwestern University
2003 Sheridan Road
Evanston, IL 60208
e-mail: barlevy@northwestern.edu

February 27, 2003

Abstract

In conventional search models where workers are allowed to sample job offers from some distribution, job mobility can play an important role in shaping the wage profile of workers. This paper attempts to quantify the role of voluntary and involuntary job mobility in accounting for the evolution of the wages of individual workers over time. Using results from record-value theory, a branch of statistics that deals with the timing and magnitude of extreme values in seuqences of random variables, I show how we can use wage data to pin down the distribution from which workers search. Applying this insight to wage data in the NLSY dataset, I show that the wage losses of displaced workers can be largely explained in terms of the foregone gains from previous job search, leaving little room for match-specific human capital.

Key Words: Search Capital, On-the-Job Search, Record Statistics

*This paper represents a much-revised version of my previous paper "Is there a Cleansing Effect of Recessions: Evidence from Wage Data." I would like to thank Joe Altonji, Marco Bassetto, Kim-Sau Chung, Chuck Manski, Francesca Molinari, Chris Taber, Nick Williams, and especially H. N. Nagaraja for helpful discussions. I also benefitted from comments at seminars in Arizona State and the Federal Reserve Bank of Minneapolis.

Introduction

Search theory has long maintained that a key challenge for workers is to secure a good job among those jobs potentially available to them. That is, if otherwise similar jobs pay different wages to the same worker, then workers will have to actively seek out those jobs that offer them particularly high wages. Various explanations have been cited as to why workers might be rewarded more on some jobs than on others, such as the notion that workers are simply better suited to certain jobs, or the notion that some frictions in the labor market (including search frictions themselves) lead firms to pursue different wage policies. But regardless of the precise reason for why workers earn different wages on different jobs, once we accept this presumption, it naturally follows that job mobility can play an important role in shaping individual wage profiles. On the one hand, if workers can search for better matches than those they already occupy, they will be able to achieve more rapid wage growth by moving to better jobs as those come along. At the same time, if recall of past offers is not possible, involuntary mobility forces workers who already secured a high-paying job to start over at a lower average wage than on the job they just lost. The purpose of this paper is to quantify the role of mobility in shaping the wage profiles of workers over time.

Conceptually, we can think of workers searching from a distribution of wages as accumulating "search capital" by securing consecutively higher-wage jobs. Analogously, if workers cannot recall previous offers, involuntary job loss will force workers to resume searching from scratch, effectively losing any search capital they previously accumulated. This interpretation is useful for distinguishing between job search and theories that emphasize match-specific human capital, i.e. skills that workers can acquire which are specific to their current job and are of no value to rival employers. Both search capital and match-specific capital imply that workers can earn higher wages by accumulating more capital, and both imply that involuntary job loss will cause them to lose any previously accumulated capital when they start their next job. But whereas search capital is accumulated by changing jobs, match-specific human capital can only be accumulated by remaining on the same job. For policy analysis as well as for predicting the effects of various shocks on labor markets, it will be important to distinguish between these two types of capital as sources of wage growth.[1]

While labor economists have devoted a great deal of effort to quantifying the role of match-specific human capital in overall wage growth, there is relatively little work quantifying the role of

[1]Workers could also accumulate general human capital, i.e. skills that are valued on all jobs, which will not be lost if a worker is forced to leave his current job. My analysis below will allow for this possibility as well.

search capital. Instead, most of the recent empirical work on search models has focused on their ability to explain the cross-sectional distribution of wages at a given point in time.[2] By contrast, this paper considers the ability of search models to account for wage *dynamics*.

To be more precise, this paper builds on the observation that the contribution of search capital to wage growth depends on the shape of the distribution of potential wage offers. That is, the shape of the distribution workers sample from uniquely determines how much a worker will gain or lose on average when he changes jobs. Gauging the potential of job mobility to affect wage growth thus reduces to estimating the underlying offer distribution. In principle, we could recover this distribution from the wages of workers on their very first job. However, once we realistically allow for unobserved heterogeneity and earnings shocks, this approach would yield a convolution of the true offer distribution and the distribution of these unobservable terms. Nevertheless, building on results in record-value theory, a branch of statistics that deals with the timing and magnitude of extreme values in sequences of random variables, I show that we can potentially recover the offer distribution even without observing the component of the wage over which workers search. Record-value theory has been previously used by statisticians to study, among other things, record temperatures, record athletic performances, road congestion, and ruin probabilities.[3] But there is also an obvious connection between record theory and job search: if workers always move to jobs that offer higher wages, their wage at any point in time will be the record highest offer since their last involuntary job change. This observation offers a way to relate observed wages back to the underlying offer distribution they were sampled from.

Using data from the National Longitudinal Survey of Youth (NLSY), I argue that wages are consistent with workers searching from a distribution of log wage offers that is exponential, i.e. from a distribution of wage offers that is Pareto. I then show that the amount of search capital workers are likely to have accumulated given their work history can account quite well for their wage losses when displaced, while match-specific human capital explains a relatively small share of this loss. Thus, search considerations on their own (together with general human capital) can reasonably explain the wage dynamics of men early on in their careers.

Turning to related work, my paper is most closely related to Topel and Ward (1992), who also investigate the role of mobility in contributing to the wage growth of young men. But their

[2] Mortensen (2002) offers a recent survey of this literature.

[3] An entertaining survey on the various applications of record statistics is provided in Glick (1978).

approach is more descriptive. In particular, they estimate the fraction of *actual* wage growth that can be explained by job mobility, as opposed to estimating an offer distribution. Moreover, their estimates combine voluntary and involuntary mobility. More recently, Keith and McWilliams (1995) examine the relationship between mobility and wages and distinguish between voluntary and involuntary turnover. However, my model suggests the relevant measure of cumulative search is the number of voluntary job changes *since the last involuntary job change*, not the number of voluntary job changes over a worker's entire history as they assume. Also related is Flinn (1986), who estimates a search model based on Jovanovic (1979). Below, I impose similar assumptions on wages as in his paper. But the nature of search in my model is different from his. First, Flinn assumes workers can only observe the quality of their match with noise, while I assume they observe it perfectly. More importantly, he assumes workers must quit to generate a new offer, while I assume workers can search on the job. Finally, my paper is related to work on returns to job seniority such as Altonji and Shakotko (1987), Abraham and Farber (1987), Topel (1991), and Altonji and Williams (1997, 1998). Since both search capital and match-specific human capital contribute to wage growth, estimating the role of one is complementary to estimating the role of the other. However, since previous work focuses on more established workers who presumably have more incentive to invest in match-specific human capital than the mobile young workers in my sample, my estimates for the returns to tenure may not be directly comparable to theirs.

The paper is organized as follows. Section 1 presents an empirical model of job search that can be used to formalize the notion of search capital. Section 2 develops the connection of the model to record-value theory and lays out the methodology for identification. Section 3 describes the data. Section 4 reports the empirical results. Section 5 concludes.

1. A Model of Job Search

I begin with a theoretical framework that formalizes the ideas set forth in the Introduction. The framework is deliberately designed to yield the familiar semi-log wage equation that has been frequently used in previous work. In some of the work I cite below, this equation is simply imposed as a primitive assumption. However, generating it explicitly in a search model will allow me to eventually derive restrictions on the unobserved match-specific component in this equation.

Standard models of search typically assume that each job offer specifies a constant wage for the duration of the match. In such an environment, a worker will naturally prefer a job if it offers him a higher wage than on his previous job. While this assumption is convenient for theoretical analysis, it needs to be modified to have any hope of according with empirical observations on wage data. For example, wages (both nominal and real) frequently change over the life of a job. In addition, a non-negligible fraction of workers who change jobs voluntarily accept offers that pay a lower wage than they previously earned. Building on Flinn (1986), I now describe a model that can accommodate these observations.

Let w_{ijt} denote the wage of individual i on job j at date t. Job j can denote any arbitrary job, not just the one the worker is employed on. When necessary, I will use the notation $j(t)$ to refer to the job the worker is employed on at date t. Flinn (1986) assumes w_{ijt} can be decomposed into two distinct terms: one that is specific to a worker-employer pair, and one that does not depend on the employer j. More precisely, the wage is multiplicatively separable in two terms, i.e.

$$w_{ijt} = \Theta_{ijt} \cdot \Phi_{it} \tag{1.1}$$

Flinn further imposes the following restrictions:

$$\Theta_{ijt} = \exp(\theta_{ij}) \tag{1.2}$$

$$\Phi_{it} = \exp(\phi_i + \beta X_{it} + \varepsilon_{it}) \tag{1.3}$$

The first assumption, (1.2), states that the match-specific term Θ_{ijt} does not vary over the duration of the match. The virtue of this specification is that it retains the main features of traditional search models without requiring that the wage be fixed on each job. In particular, since changes in Φ_{it} affect wages on all jobs in the same way, the worker will prefer the employer with the highest θ_{ij}, just as in the traditional model he prefers the employer offering the highest constant wage. An additional virtue of this specification is that it allows for the possibility that workers voluntarily move to jobs that pay less than what they earned in their previous jobs. This would occur if the time-varying term Φ_{it} fell at the same time a worker happens to switch to a match with a higher θ_{ij}. Although the worker earns a lower wage on his new job than he previously earned at his old job, he would have earned an even lower wage had he stayed on his original job.

For each worker i, we can summarize the distribution of θ_{ij} across potential employers j by a cdf $F_i(\cdot)$. I assume this distribution is continuous and identical across workers i, i.e. $F_i(\cdot) = F(\cdot)$ for all i. I will refer to θ_{ij} as match quality, although one could equally interpret differences in θ_{ij} across employers as heterogeneity in wage policies, in line with Burdett and Mortensen (1998).

Note that the assumption that Θ_{ijt} is time-invariant denies a role for match-specific human capital, since any wage growth on one job must automatically carry over to other jobs as well. So as not to rule out match-specific human capital from the outset, I generalize this setup to allow Θ_{ijt} to depend on the amount of time T_{ijt} worker i has spent working on job j by date t. In particular, suppose there exists some constant $\gamma \geq 0$ such that

$$\Theta_{ijt} = \exp\left(\theta_{ij} + \gamma T_{ijt}\right) \tag{1.4}$$

where once again $\theta_{ij} \sim F(\cdot)$. Under (1.4), at each date t it will be optimal for the worker to choose the job offering the highest Θ_{ijt}. This is because wages grow at the same rate on all jobs, so that if $\Theta_{ijt} > \Theta_{ij't}$ at date t for two jobs $j \neq j'$, the same inequality will hold at any date beyond t. This result hinges on the assumption that returns to tenure in (1.4) are linear. For more general specifications, a job that offers a higher wage than some other job at one point in time might offer a lower wage at another point in time, and we would need to know the rate of time preference, interest rate, and utility function of the worker to determine which job is preferable.[4]

Assumption (1.3) states that the match-independent term Φ_{it} consists of three components: ϕ_i, which captures fixed worker attributes such as schooling or ability; X_{it}, which captures time-varying worker attributes, i.e. labor market experience; and ε_{it}, which denotes a stochastic earnings shock but can also encompass measurement error. Flinn (1986) assumes ε_{it} is i.i.d., while Topel and Ward (1992) argue it is $\Delta\varepsilon_{it}$ that is i.i.d.[5] As will be clear below, my approach only requires that $E\left(\Delta\varepsilon_{it}\right) = 0$. Substituting in (1.3) and (1.4) and taking logs yields the familiar semi-log wage regression that has been used to model wages in previous work, i.e.

$$\ln w_{ijt} = \beta X_{it} + \gamma T_{ijt} + \phi_i + \theta_{ij} + \varepsilon_{it} \tag{1.5}$$

As noted earlier, the virtue of deriving this wage equation explicitly in a search setting is that this will allow us to derive explicit restrictions on θ_{ij}, as will be shown below.

Next, I describe the assumptions that govern how workers choose jobs. At the beginning of each period, an employed worker is forced to leave his job with probability $s \in (0, 1)$. Otherwise, he can stay on his current job. That same period, with probability $p \in (0, 1)$ the worker learns of a

[4] Specifically, when $\Theta_{ijt} = \exp\left(\theta_{ij} + g\left(T_{ijt}\right)\right)$ for some increasing function $g(\cdot)$, a worker with a match of quality θ and tenure T will set a reservation cutoff $\theta^*(\theta, T)$ and accept a new match if and only if its match quality exceeds θ^*. If $g(\cdot)$ is linear, $\theta^*(\theta, T) = \theta + g(T)$, regardless of the discount rate and interest rate. When $g(\cdot)$ is not linear, though, it will often be impossible to obtain a closed form expression for θ^*.

[5] Topel and Ward estimate the autocovariance of the residual in the regression $\Delta \ln w = H(X, T) + \Delta\varepsilon$ and argue ε is approximately a random walk with drift. They note that "about 95% of the within-job residual variance in measured quarterly wages is associated with the systematic component; the remainder is accounted for by measurement error or other transitory shocks to measured earnings." Abowd and Card (1989) obtain similar results.

new job, whose match quality θ is drawn from $F(\cdot)$. The worker observes θ on the new job — in contrast to Flinn's model where he can only observe the sum $\theta + \varepsilon$ — and chooses whether to stay on his current job (alternatively, unemployed) or accept the new offer. Once a job is declined, it cannot be recalled. This specification is consistent with conventional assumptions in the literature, e.g. Pissarides (2000). I will classify a job change that occurred because the worker had to leave his employer as involuntary, and a job change that was initiated by the worker as voluntary.

An important feature of the search technology above is that workers can receive offers while employed, i.e. they can engage in on-the-job search. This feature gives rise to the notion of search capital introduced earlier. Formally, define the stock of search capital for a given worker as the quality of his current job $\theta_{ij(t)}$. For an unemployed worker, this value will be set to the minimum value in the support of $F(\cdot)$. As specified, the model implies that each voluntary job change increases the stock of search capital. This is because as long as returns to tenure are linear, a worker will prefer a new job j' to his previous job j if and only if $\Theta_{ij't} > \Theta_{ijt}$. Since the worker has no previous tenure on the new job j', it follows that $\theta_{ij'} \geq \theta_{ij} + \gamma T_{ijt} \geq \theta_{ij}$. Hence, each voluntary job change will add a (random) increment to the stock of search capital. If the worker is involuntarily displaced, the absence of recall implies the previous value of $\theta_{ij(t)}$ will be reset to the minimum possible value, so that any previously accumulated search capital will be lost.

In sum, the model suggests two ways a worker can increase his wages over time. The first is to increase Φ_{it}. This can be viewed as investment in general human capital, since it allows the worker to earn higher wages on *all* jobs. The second is to increase Θ_{ijt}, which can be achieved in two ways: building up tenure T_{ijt} with a given employer j, or searching through the offer distribution to arrive at an employer with a high θ_{ij}. In contrast to investment in Φ_{it}, any gains in Θ_{ijt} will be lost when a worker is involuntarily displaced. The purpose of the empirical work below is to quantify the role of these channels in accounting for the wage changes of young men.

2. Empirical Methodology

To be sure, previous work has already attempted to quantify some of the sources of wage growth discussed above. For example, Altonji and Shakotko (1987), Abraham and Farber (1987), and Topel (1991) begin with an equation of the form (1.5) and proceed to estimate the returns to experience β and seniority γ. All three recognize the importance of match quality θ_{ij}. However, rather than attempt to gauge its role in shaping wage dynamics, they treat it as a nuisance for estimating β and γ. For example, the first two papers devise instruments uncorrelated with θ_{ij} to estimate these coefficients. But they do not then use their estimates to infer the role of mobility. Topel proposes an alternative two-step method for estimating β and γ, which yields much larger estimates for γ for the same sample used in earlier work, but he too ignores the subsequent implications for the role of θ_{ij} in accounting for wage growth. By contrast, my approach is focused on estimating the role of unobserved match quality θ_{ij} directly, and treats the task of assigning values to β and γ as an intermediate step towards that goal.

As noted in the Introduction, quantifying the role of search capital essentially boils down to estimating the offer distribution $F(\cdot)$. In simpler search settings, identifying $F(\cdot)$ is trivial. For example, in a related model that abstracts from unobserved heterogeneity, Mortensen (2002) suggests using the empirical wage distribution for workers on their very first job. This approach would not work here; even with estimates of β and γ, we would only be able to isolate the sum $\theta_{ij}+\phi_i+\varepsilon_{it}$, so that we can only identify a convolution of $F(\cdot)$ and the distribution of these additional terms. Nevertheless, I show that when $\gamma = 0$, we can uniquely pin down $F(\cdot)$. Intuitively, given two distinct distributions $F(\cdot)$, the rate at which a worker accumulates search capital – i.e. the rate at which wages rise with the number of voluntary job changes – will be sufficiently different between the two that we can distinguish between them with enough data. Unfortunately, my proof cannot be easily adapted to the case of $\gamma > 0$. Still, I show that even when $\gamma > 0$, we can easily check if $F(\cdot)$ is at least consistent with particular functional forms.

In order to establish these results, I first recast the search model above as a *record model*. Building on insights from record-value theory, I then show how one can potentially identify the offer distribution $F(\cdot)$ from wage data. Lastly, I point out an overidentifying restriction of the model that can be used to further assess the validity of this estimate.

2.1. Record Statistics and Job Search

I begin with a brief introduction to record-value theory. Given a sequence of random variables, an element in the sequence is defined as a *record* if it exceeds the realized value of all observations that preceded it in the sequence. Statisticians have devised models to analyze the occurrence of records in sequences of random variables, and have used them to study various applications such as weather patterns (e.g. using the frequency of record temperatures to detect global warming). However, these models have not been utilized much in the economics literature.[6]

[6] One exception is Kortum (1997), who considers a model in which innovation involves drawing from a distribution of technologies. A successful innovation is one that is more productive than all techniques that came before it.

The first model of records was developed in Chandler (1952). He assumed an infinite sequence of random variables $\{\theta_m\}_{m=1}^{\infty}$ in which each of the individual observations is i.i.d. He then defined record times within this sequence as follows. Let $M_1 = 1$, and for any integer $n > 1$, let

$$M_n = \min\left\{j : \theta_j > \theta_{M_{n-1}}\right\}$$

The n-th record, which is denoted by R_n, is just the value of θ_m at the n-th record time, i.e. $R_n = \theta_{M_n}$. As an illustration, consider the following sequence of daily temperatures at a given location as measured on the same date each year:

$$\{65, 61, 68, 69, 63, 67, 71, ...\}$$

The first observation is trivially a record, so $M_1 = 1$ and $R_1 = 65$. The next observation that exceeds this value is the third one, so $M_2 = 3$ and $R_2 = 68$. The very next observation exceeds this value, so $M_3 = 4$, and $R_3 = 69$, and similarly $M_4 = 7$ and $R_4 = 71$. Thus, the sequence of records $\{R_n\}_{n=1}^{\infty}$is given by $\{65, 68, 69, 71, ...\}$. Note that each record can be alternatively viewed as an order statistic, i.e. R_n is the maximum of a sample of M_n draws, but where the sample size M_n is random. Thus, a record will have the same distribution as a mixture of order statistics. But since this mixture is over an infinite number of possible distributions, a record statistic may not inherit the properties of its corresponding order statistics. For example, as demonstrated in Nagaraja (1978), the moments of a record may not exist even when the moments of the associated order statistics exist for any finite sample size. For this reason, record statistics have emerged as a distinct but parallel branch of extreme-value theory from order statistics.

For the classical record model above, statisticians have been able to derive the distribution of record times M_n, the distribution of the number of records within a given sample size, the distribution of record values R_n for a given parent distribution $F(\cdot)$, and the asymptotic distribution of record values R_n (if it exists). More recently, statisticians have tackled more complicated record processes, e.g. when observations are no longer i.i.d. For excellent surveys of these results, see Arnold, Balakrishnan, and Nagaraja (1992, 1998) and Nevzorov and Balakrishnan (1998).

It is easy to see why record-value theory might be useful for analyzing search models. Empirically, we often collect data only on jobs a worker accepts. Hence, in contrast with temperature data where

---

Kortum remarks on the connection between his model and record theory. However, most of his analysis uses results from order statistics rather than record statistics, since he essentially conditions on sample size rather than on the number of previously successful innovations. Another exception is Munasinghe, O'Flaherty, and Danninger (2001), who analyze the number of track and field records in national and international competitions to study the effects of globalization, and remark on the likely applicability of record theory in economics, including job search.

we can measure the temperature each year regardless of its value, wage data is only collected if the present wage offer is more attractive to the worker than all previous offers. In other words, available wage data only corresponds to record-setting values $\{R_n\}_{n=1}^{\infty}$, not the sequence of observations $\{\theta_m\}_{m=1}^{\infty}$. But by explicitly recognizing this record structure, we might still be able to relate wage data back to the offer distribution from which workers draw offers. Conversely, given an offer distribution, we can compute how much a worker should gain on average as he accumulates search capital using the appropriate record statistics for repeated draws out of this distribution.

2.2. Identifying the Offer Distribution

I now examine whether the record structure implicit in the search model can be used to identify $F(\cdot)$. Since workers cannot recall offers they previously turned down, the search process is reset each time a worker is forced to leave his job. Thus, let n denote the number of jobs a worker has held *since his last involuntary job change*, or else since he first entered the labor market. Let $\theta_{it} \equiv \theta_{i,j(t)}$ denote the quality of the match on job $j(t)$, and similarly $T_{it} \equiv T_{i,j(t),t}$ as the time the worker spent on that job by date t. Finally, let θ_{it}^n reflect the fact that $j(t)$ is the worker's n-th job. First-differencing (1.5) to get rid of unobserved heterogeneity, and using the fact that $\Delta X_{it} = T_{it} = 1$ if a worker remains with the same employer, we obtain

$$\Delta \ln w_{it} = \begin{cases} \beta + \gamma + \Delta\varepsilon_{it} & \text{if same job} \\ \beta - \gamma T_{i,t-1} + \left(\theta_{it}^{n+1} - \theta_{i,t-1}^{n}\right) + \Delta\varepsilon_{it} & \text{if voluntary change} \\ \beta - \gamma T_{i,t-1} + \left(\theta_{it}^{1} - \theta_{i,t-1}^{n}\right) + \Delta\varepsilon_{it} & \text{if involuntary change} \end{cases} \tag{2.1}$$

I begin with the case where $\gamma = 0$, i.e. where there are no returns to seniority. Let M denote the (random) number of offers the worker receives between consecutive involuntary job changes, and let $m = 1, 2, ..., M$ index individual offers. Define θ_m as the value of θ on the m-th offer. Since the worker will only change jobs if he encounters an offer with higher match quality, it follows that $\theta_{it}^n \overset{d}{=} R_n$, i.e. the match quality on the worker's n-th job will have the same distribution as the n-th record from a sequence of i.i.d. draws from $F(\cdot)$. The wage changes of workers who change jobs, either voluntarily or involuntarily, thus contain a difference between record values.

Remark 1: The search model above differs from the classical record model in that rather than observing an infinite sequence $\{\theta_m\}_{m=1}^{\infty}$, we get to observe a sequence $\{\theta_m\}_{m=1}^{M}$ of random length.

In the Appendix, I show that M has a geometric distribution, i.e.

$$\begin{aligned}\Pr(M=m) &= \left(\frac{p(1-s)}{p+(1-p)s}\right)^{m-1}\frac{s}{p+(1-p)s} \\ &\equiv (1-\alpha)^{m-1}\alpha\end{aligned}$$

The number of offers a worker receives before being laid off is therefore finite with probability 1, in contrast to the classical model where the number of observations is infinite. Bunge and Nagaraja (1991) already analyzed several implications of this setup. As they point out, since we may not always observe n records in our sample, the distribution of record values is implicitly conditioned on observing that many records. That is, if N denotes the number of records in $\{\theta_m\}_{m=1}^{M}$, any analysis of the n-th record is conditioned on the event that $N \geq n$. Bunge and Nagaraja (1991) show that if M is geometric, N will be a truncated Poisson, i.e. $\Pr(N=n) = \frac{\alpha(-\ln\alpha)^{n+1}}{(1-\alpha)n!}$, and proceed to derive an explicit expression for the joint distribution of the first n records conditional on $N \geq n$ for any parent distribution $F(\cdot)$. ■

The observations above allow us to rewrite equation (2.1) for voluntary job changers as

$$\Delta \ln w_{it} = \beta \Delta X_{it} + (R_{n+1} - R_n) + \Delta\varepsilon_{it} \tag{2.2}$$

Taking expectations conditional on observing at least $n+1$ records, we have

$$E(\Delta \ln w_{it} - \beta\Delta X_{it} \mid N > n) = E(R_{n+1} - R_n \mid N > n) \tag{2.3}$$

Note that from (2.1), we can identify the returns to experience β off of the wage gains of job stayers. Thus, we can recover the conditional expectations $E(R_{n+1} - R_n \mid N > n)$. I now argue that these moments uniquely establish the shape of the parent distribution $F(\cdot)$. That is, by observing the average wage gain among all workers who move from their first job to their second job, from their second job to their third job, and so on, we can uniquely identify the shape of the offer distribution $F(\cdot)$. First, though, I need the following lemma:

Lemma: Consider a sequence $\{\theta_m\}_{m=1}^{M}$ where $M \sim \text{geo}(\alpha)$. Let $\{R_n\}_{n=1}^{N}$ denote the records of this sequence. If $E(|\theta_m|) < \infty$, then $E(R_{n+1} - R_n \mid N > n) < \infty$ for $n = 1, 2, 3, ...$

Armed with this lemma, I establish the following result:

Proposition:[7] Consider a sequence $\{\theta_m\}_{m=1}^{M}$ where $\theta_m \sim F(\cdot)$ such that $E(|\theta_m|) < \infty$. If

[7] I am indebted to H. N. Nagaraja for suggesting the proof of this Proposition.

M has a geometric distribution, the sequence $\{E(R_{n+1} - R_n \mid N > n)\}_{n=1}^{\infty}$ uniquely characterizes the distribution $F(\cdot)$ in the set of continuous distribution functions, up to a location parameter.

Remark 2: A related result for the case where $\Pr(M = \infty) = 1$ was established by Gupta (1984). He showed that under additional regularity conditions, the sequence $\{E(R_{n+1} - R_n)\}_{n=1}^{\infty}$ exists and uniquely characterizes $F(\cdot)$ in the set of continuous distribution functions, up to a location parameter. However, the proposition above involves conditional moments, since we are not always assured of reaching the n-th record, and thus requires a different method of proof than in Gupta. Interestingly, it turns out that the characterization result in the geometric random record model is more general than in the classical record model, i.e. it is also true for subsequences of record spacings whose indices satisfy the Müntz-Szász criterion, something that is not true in the classical record model. For details, see Nagaraja and Barlevy (2002). ■

Remark 3: The proposition above ensures that the sequence $\{E(R_{n+1} - R_n \mid N > n)\}_{n=1}^{\infty}$ maps to a unique continuous distribution $F(\cdot)$, up to a location parameter. However, its proof does not offer a way to readily construct $F(\cdot)$. In practice, I proceed in the opposite direction: I begin with a candidate distribution $F(\cdot)$, and then check whether the expected record spacings for this distribution are consistent with the empirical moments estimated from the data. To the extent that the two are consistent, we can use the proposition above to conclude that the offer distribution must take the same shape as our candidate distribution, since any other continuous distribution would have necessarily failed this test (at least given enough data). ■

The proposition above establishes that we can uncover the distribution of θ from observations on $\Delta\theta_{it} + \Delta\varepsilon_{it}$. The reason we can overcome the confounding $\Delta\varepsilon_{it}$ term is that we can assign each observation to a record number n. Since the value of n is identified from job mobility rather than from $\Delta\theta_{it} + \Delta\varepsilon_{it}$, it is related to $\Delta\theta_{it}$ but not to $\Delta\varepsilon_{it}$. This additional information is enough to allow us to isolate the distribution of θ, even though we cannot observe $\Delta\theta$ directly.

As noted above, my empirical implementation involves checking whether the data is consistent with a candidate functional form for $F(\cdot)$. One of the distributions that figures prominently in record-value theory, and is thus an obvious candidate, is the exponential distribution

$$F(x) = 1 - e^{-x/\lambda}$$

for some $\lambda > 0$. In this case, the average record margin is independent of n, i.e.

$$E(R_{n+1} - R_n \mid N > n) = \lambda$$

for $n = 1, 2, 3, ...$ This result is due to the memoryless property of the exponential distribution, which states that if a random variable X is exponential, $\Pr(X > x + y \mid X > y) = \Pr(X > x)$ for all $x, y \geq 0$. Hence, to check if $F(\cdot)$ is exponential, we only need to examine whether the average wage gain for voluntary job changers net of returns to experience varies with n. What makes the exponential case especially convenient is that we don't even need to estimate the parameter $\alpha = s(p + (1-p)s)^{-1}$ to carry out this test.[8] But the exponential case is not just convenient; it is also of *a priori* empirical relevance. In particular, if θ is exponential, the distribution of wages from which workers search will be Pareto. The Pareto distribution arises repeatedly in empirical distributions of earnings, and so is a natural one to consider here as well.[9]

Next, consider the case of positive (but linear) returns to tenure, i.e. $\gamma > 0$. Once again, we can relate the wage gains of voluntary job changers to record moments. In particular, let t_m denote the number of periods between when the worker started working on his first job (for which $n = 1$) and when he received his m-th offer. Define a new variable $\widetilde{\theta}_m = \theta_m - \gamma t_m$. The distribution of $\widetilde{\theta}_m$ is a mixture of affine translations of $F(\cdot)$, i.e.

$$\Pr(\widetilde{\theta}_m < x) = \sum_{s=m-1}^{\infty} \Pr(t_m = s) \cdot F(x + \gamma s)$$

Recall that the worker will only switch to a new job if the difference in match quality between the new job and his old job exceeds the cumulative returns to tenure on his old job. This is equivalent to the condition that the current value $\widetilde{\theta}_m$ exceeds all past realizations $\widetilde{\theta}_1, ..., \widetilde{\theta}_{m-1}$. Thus, just as in the case of $\gamma = 0$, a worker will change jobs if and only if he encounters a record value within an appropriately defined sequence of random variables. If we define $\widetilde{R}_n$ as the value of the n-th record in the sequence $\{\widetilde{\theta}_m\}_{m=1}^{M}$, a little algebra reveals that the wage gain of voluntary job changers net of returns to experience is now equal to

$$E(\Delta \ln w_{it} - \beta \Delta X_{it} \mid N > n) = E(\widetilde{R}_{n+1} - \widetilde{R}_n \mid N > n) + \gamma \tag{2.4}$$

Once again, then, we can use wage data to construct conditional moments from a record process. In contrast to the case where $\gamma = 0$, the individual observations $\widetilde{\theta}_m$ are no longer independent

[8]Even though we don't need to estimate α in this case, we still implicitly use the assumption that α is the same for all agents. This is because I have not established that the proposition above extends to the case where α varies across agents, i.e. I have not shown that record moments uniquely identify $F(\cdot)$ when M is a mixture of geometrics. Even if this characterization result is not true, we could still in principle disaggregate the data by α and identify the distribution from individual subgroups.

[9]Neal and Rosen (2000) discuss the prevalence of the Pareto distribution in earnings data. Mortensen (2002) cites work using Danish data that specifically seeks to estimate a wage *offer* distribution, based on a different approach than the one presented here, and likewise reports evidence of a Pareto tail.

nor identically distributed. However, the distribution of each $\widetilde{\theta}_m$ can still be summarized in terms of a single parent distribution $F(\cdot)$, so there is some hope that record moments might be useful for recovering the distribution of interest.[10] Unfortunately, the argument for the i.i.d. case cannot be adapted to this more general case, and thus we are no longer assured that the sequence $\{E(\widetilde{R}_{n+1} - \widetilde{R}_n \mid N > n)\}_{n=1}^{\infty}$ can pin down the shape of the offer distribution $F(\cdot)$.[11]

Despite this complication, there is still a simple way to check the particular hypothesis that $F(\cdot)$ is exponential. In fact, we can proceed in exactly the same way as when $\gamma = 0$. Recalling that $\Delta X_{it} = \Delta T_{it} = 1$, a regression of $\Delta \ln w_{ijt}$ on ΔX_{it} for job stayers would yield a coefficient for β that combines returns experience and returns to seniority, i.e. $\widehat{\beta} = \beta + \gamma$. At the same time, the memoryless property of the exponential distribution implies that

$$E(\widetilde{R}_{n+1} - \widetilde{R}_n \mid N > n) = \lambda$$

Thus, if $F(\cdot)$ is truly exponential, the average wage gains of voluntary job changers net of the (incorrectly) estimated returns to experience $\widehat{\beta}$ will appear not to vary with n. But again, we need to keep in mind that we have not established there isn't some other distribution for which $E(\widetilde{R}_{n+1} - \widetilde{R}_n \mid N > n)$ is also constant for all n (at least in the limit when time is continuous; as noted in the footnote above, in discrete time $F(\cdot)$ is *not* uniquely identified in the set of continuous distributions). In other words, the test involves a necessary condition for $F(\cdot)$ to be exponential, and it is not clear whether this condition could also be sufficient, at least under some additional assumptions.

2.3. Overidentifying Restrictions: the Wage Losses of Involuntary Job Changers

Assuming we can use the procedure above to obtain an estimate of $F(\cdot)$, it would be desirable to come up with an overidentifying test that can be used to further establish the validity of our

[10] In the case of non-linear returns to tenure, we can similarly define some $\widetilde{\theta}_m$ so that a worker will change jobs if and only if he encounters a new record within the sequence $\{\widetilde{\theta}_m\}_{m=1}^{M}$. In this case, not only are the individual observations not i.i.d., but the distribution of each observation depends on whether a record was set on the previous observation. This variation on the classical record model was previously explored in Pfeifer (1982).

[11] In fact, as the model is formulated, record moments *cannot* uniquely identify all continuous distributions when $\gamma > 0$. First, for distributions with bounded support, the number of records we observe is finite, so there will not be enough moments for unique identification. Even with infinitely many records, we cannot uniquely identify distributions whose support is bounded from below. Let $\underline{\theta}$ denote the infimum of the support of such a distribution. The shape of the distribution in the interval $[\underline{\theta}, \underline{\theta} + \gamma]$ will matter for $E(\widetilde{R}_1)$, but not for $E(\widetilde{R}_n \mid N \geq n)$ for $n \geq 2$. Thus, we cannot distinguish between two distributions that differ only within this interval and share the same value for $E(\widetilde{R}_1)$. These problems would go away in the limit when time is continuous; however, the continuous-time formulation is sufficiently involved that it is best left for a separate paper.

estimate. Since the above methodology uses the wage gains of voluntary job changers, a natural candidate for such a test is the wage losses of involuntary job changers. In particular, given an estimate of γ and the offer distribution $F(\cdot)$, we should be able to compute the implied average wage loss for an involuntary job changer, i.e. $E(\Delta\Theta_{ijt})$, and then compare this to the average wage loss in the data as a consistency check on our estimate.

Again, I begin with the case where $\gamma = 0$. In this case, the only reason an involuntary job changer should suffer a wage loss (net of returns to experience) is if the match quality on his new match is lower than on the job he just left. The match quality for a worker who is forced to leave his n-th job will correspond to the n-th record from a sequence of i.i.d. draws from $F(\cdot)$ with a geometric number of observations. Here, we would want to condition on the event that the total number of records N in the sequence is equal to n, since we know that the n-th job is the worker's last. Similarly, the match quality on his new job corresponds to the first record from such a sequence, conditional on $N \geq 1$. But this condition is trivially true. Thus, we have

$$E\left(\, |\Delta \ln w_{it} - \beta \Delta X_{it}| \;\middle|\; N_{t-1} = n \right) = E(R_n \mid N = n) - E(R_1) \tag{2.5}$$

By an extension of the analysis in Nagaraja and Barlevy (2002), we can establish that the sequence $\{E(R_n \mid N = n) - E(R_1)\}_{n=1}^{\infty}$ uniquely characterizes $F(\cdot)$ within the set of continuous distributions, up to a location parameter. Thus, these wage losses constitute a legitimate overidentifying test on $F(\cdot)$ when $\gamma = 0$. However, these moments generally depend on the mobility parameter α, even when $F(\cdot)$ is exponential. To avoid having to also estimate α, I follow an alternative route of deriving a lower bound for the expression on the right-hand side of (2.5) that is independent of α. This allows me to proceed without having to separately identify α. But the disadvantage of this approach is that it yields a weaker test that can reject only certain alternative distribution functions $F(\cdot)$.

To construct this bound, let $\theta_{m:m}$ denote the maximum of the first m elements of a sequence of random variables. By conditioning on the number of offers M between consecutive separations, we obtain

$$E(R_n \mid N = n) = \sum_{m=n}^{\infty} \omega_{mn} E(\theta_{m:m} \mid M = m,\ N = n)$$

where $\omega_{mn} = \Pr(M = m \mid N = n)$ denotes the fraction of all workers involuntarily separated from their n-th job who received exactly m offers before separation. Since for any $m \geq n$, we have

$$E(\theta_{m:m} \mid M = m,\ N = n) = E(\theta_{m:m}) \geq E(\theta_{n:n})$$

It then follows that $E(R_n \mid N = n) \geq E(\theta_{n:n})$. Thus, a lower bound on the average wage loss net of the returns to experience is given by $E(\theta_{n:n} - \theta_{1:1})$. If $\theta \sim \exp(1/\lambda)$, it is well-known that

$$E(\theta_{n:n}) = \lambda \sum_{i=1}^{n} \frac{1}{i} \equiv \lambda g(n)$$

Thus, if θ is distributed as an exponential, we should see

$$E\left(|\Delta \ln w_{it} - \beta \Delta X_{it}| \;\middle|\; N_{t-1} = n \right) \geq \lambda (g(n) - 1) \tag{2.6}$$

where λ is estimated from voluntary job changers. This inequality will be the overidentifying restriction I will focus on in my empirical analysis.

Similarly, one could compute the expected wage loss for a displaced worker when $\gamma > 0$. In this case, the average wage loss can be expressed in terms of moments of the record process $\widetilde{R}_n$ defined above. However, as I argue below, returns to tenure in my sample of young workers appear to be quite small, so that the case of $\gamma = 0$ should provide a good first approximation. To preview my results, the average wage losses of involuntary job changers in the data are indeed quite close to the lower bound in (2.6) when I use the estimate of λ from voluntary job changers. This finding suggests that $\omega_{nn} \approx 1$, since this is when the bound in (2.6) is exactly binding. At first glance, it might seem surprising that most workers on their n-th job appear to have received exactly n offers. However, there are several explanations that can account for this result. First, a relatively high separation rate s, as is true among young workers, would tend to shift the steady-state distribution towards lower values of m for a given n. Intuitively, a higher separation rate makes it more difficult to observe a worker that has accumulated too many offers, since they are likely to be displaced before they accumulate enough such offers. In addition, the fact that the sample covers the first few years in the labor market biases the distribution towards putting more weight on lower values of m for a given n, simply because workers have not been in the labor market for enough time to accumulate many offers.

Another explanation, which is specific to the exponential distribution, involves the fact that for a given m, the difference $g(m) - g(n) = \sum_{i=n+1}^{m} \frac{1}{i}$ will be quite small for large n, at least when m is not much larger than n. Thus, $\omega_{nn} = 1$ might serve as a good approximation in some cases even if $\omega_{mn} > 0$. Finally, suppose s and p vary over the duration of the match. In particular, suppose s starts high at the beginning of a match but declines over time, while p starts out low but rises over time. This would tend to minimize the odds of observing a worker who accumulated many offers before he was separated, particularly for lower values of n. Both assumptions seem plausible. On the one hand, various jobs involve an initial probation period when the worker is more likely to be laid off. At the same time, job-shopping considerations such as Jovanovic (1979) suggest workers are more likely to leave for a new match the longer they have been employed, since they have more precise information about the quality of the match and thus have a lower option value of staying with a job. However, under these assumptions, the number of offers M a worker receives before being separated will not be distributed as a geometric, and the proposition above may not apply. In particular, an exponential $F(\cdot)$ would still imply that average wage gains should not vary with n, but it might no longer be the case that this is the unique distribution for which this is true.

3. Data

To implement the analysis above, I need a dataset with detailed work-history data that can be used to construct n. In addition, since job mobility is highest when workers first enter the labor market, it seems wise to focus on young workers; this both maximizes the sample of job changers, and suggests the incentive to invest in match-specific human capital is low. These considerations suggest using the National Longitudinal Survey of Youth (NLSY) dataset. The NLSY follows a single cohort of individuals who were between 14 and 21 years old in 1979. At the time I assembled my dataset, data was available through 1993. In each year, respondents were asked questions about all jobs they worked on since their previous interview, including starting and stopping dates, the wage paid, and the reason for leaving the job. To mitigate the influence of mobility due to non-wage considerations, e.g. pregnancy or child-care, I restrict attention to male workers.

With the exception of n, my variables are standard. For the wage, I use the hourly wage as reported by the worker for each job, divided by the GDP deflator (with base year 1992). I also experimented with using the CPI, but the results were similar. To minimize the effect of extreme outliers, which are likely to be due to coding errors, I trimmed observations for which the reported wage was less than or equal to \$0.10 or greater than or equal to \$1000. This eliminated 0.1% of all wage observations. To construct a measure of potential experience, I dated each worker's entry into the labor market at his birthyear plus his reported years of schooling (highest grade completed) plus 6. However, if an individual reported working on a job prior to that year, I dated entry at the year in which he reports his first job. This raises the concern that my results could be affected by summer or after-school jobs. But such jobs are only a small part of my sample, and, for reasons explained below, these jobs are unlikely to enter into my calculations of the conditional record moments $E(R_{n+1} - R_n \mid N > n)$. Table 1 provides summary statistics for my sample.

The one novel variable in my analysis is the number of jobs n a worker has been employed on since his last involuntary job change, and so it merits some discussion. The determination of whether the worker left his job voluntarily or involuntarily is based on the reason cited by the worker for leaving his job. For jobs that stretch out over multiple interviews, workers will sometimes offer different reasons for leaving the same job in different surveys. In these cases, I used the explanation cited in the last year the worker reported being on the job. Although the precise wording of the question varies from year to year, roughly speaking, a voluntary job change corresponds to a self-reported quit while an involuntary job change corresponds to a layoff, discharge, or end of a job. This relies heavily on the accuracy of the reported reasons. However, one can view the empirical work as a way to validate the accuracy of these self-reports, since the wage gains of voluntary job changers and the wage losses of involuntary job changers are likely to be inconsistent if the reason for job change are systematically misreported.

Several complications arise in constructing n from work-history data. First, a non-trivial fraction of workers in the NLSY hold more than one job at a point in time, which raises the question of how to deal with overlapping jobs. In this regard, I use the analysis in Paxson and Sicherman (1996) that argues the primary reason workers hold multiple jobs is that they are constrained to work a maximum number of hours on each job. Suppose workers are equally constrained on all jobs and can work on only one job full time. However, workers can receive additional draws from the distribution $F(\cdot)$ and work on those jobs on a part time basis. If we observe a worker already employed in job A who takes on a second job B, two outcomes are possible. First, the worker might leave job B and continue to work on job A. If secondary jobs are only available on a part-time basis, this does not provide us with any information on match quality on the worker's primary job. I therefore ignore job B in my analysis, i.e. it is as if the worker had never taken on a secondary job. The other possibility is that the worker leaves job A and remains in job B. This implies an opening for a full-time position must have opened up at some point on job B. Since I assume secondary jobs are drawn from the same distribution $F(\cdot)$, job B is just another draw from $F(\cdot)$, and so I treat it the same way as a new job that starts only after job A ends.

Out of the 52,827 jobs in my original sample, the procedure above identifies 8,232 as secondary jobs. To check if this classification is sensible, the NLSY asks workers to rank their jobs each year in terms of which is their primary job. Of the 8,232 jobs I identify as secondary jobs, 72% are never ranked as the primary job in any year, and only 9% are identified as the primary job in each year the job is reported, most of which involve only a single observation.

After eliminating secondary jobs, I order remaining jobs according to their starting dates. Starting with the first time a worker reports an involuntary job change, I assign his next job a value of $n = 1$. If the worker reports leaving his n-th job voluntarily, the next job is assigned a value of $n + 1$. If the worker reports leaving his job involuntarily, I reset this number to 1. If no reason is offered for leaving the job, I must wait until the next involuntary job to reassign n. Out of the 44,595 jobs that remain after eliminating secondary jobs, I was able to assign a value of n for 26,929, or 60%. In an attempt to raise this number, I also tried to identify entry into the labor force and assign $n = 1$ to the first job the worker accepts to overcome the left-censoring problem. However, the number of observations identified this way was small, and adding it to my sample had no effect on my results. Thus, I only report results for the case where n is assigned after an involuntary job change. As alluded to above, this has the advantage that it is more likely to exclude summer and after-school jobs, since we assign n only after we observe at least one involuntary job change. While a significant fraction of my observations could not be assigned a value of n, recall from equation (2.1) that the wage changes of job stayers do not depend on n, so I can still use some of the data for which no value of n could be assigned.

To conform with the timing of my model, I chose one job per worker each year. To be included in the sample, the worker must have reported working on that job within two weeks of the interview date. This insures that observations are gathered on a regular basis, in accordance with my timing assumptions. The choice of an annual horizon is merited by the finding in Topel and Ward (1992) that there is a "strong tendency for within-job earnings changes to occur at annual intervals." In cases where the worker is employed on more than one job at the time of the interview even after eliminating secondary jobs (i.e. he has yet to move full time to his new job), I use the job that began first. Thus, a job is identified as the primary job only in the first year the worker is observed to work exclusively on that job. Proceeding this way yields a sample of 56,962 observations on 25,424 jobs for 6,139 workers. Since first-differencing requires me to use at least one lag, and since not all of the observations have the requisite data on wages, education, or dates of employment, my final sample is a little more than half the size of the original sample.

In Figure 1, I report the distribution of n for all 56,962 observations in Figure 1. Figure 1a shows the fraction of jobs for which I could not assign a value to n. This value falls from 84% in 1979 to 29% in 1993. This ratio stabilizes at this level, reflecting the fact that a non-negligible fraction of workers do not report an involuntary job loss over the duration of the sample. Figure 1b shows the distribution of n for those jobs where n could be assigned. The graph suggests the distribution of n settles down after 10 years, with about 40% of all observations associated with $n = 1$, 25% with $n = 2$, 15% with $n = 3$, and virtually all remaining observations with $n \leq 7$.

The average value of n for all observations where n could be assigned rises from 1.2 in 1979 to 2.4 by 1993. Note that few observations involve a very large value for n. This is consistent with a standard result in record-value theory that records from a sequence of i.i.d. draws should be relatively rare (and even rarer if $\gamma > 0$).

The fact that few workers have a high value of n underscores the importance of distinguishing between voluntary and involuntary job changes. Topel and Ward (1992) look at cumulative mobility, and estimate that 96% of workers were employed on more than one full-time job after 10 years of potential experience, and half held more than 6 jobs. If these job changes were all voluntary, this would suggest significant cumulative search experience. But Figure 1b reveals that when we take into account that some job changes are involuntary, even after 10 years, a significant fraction of workers should essentially be viewed as being on their first job.

4. Empirical Analysis

I now proceed to use the data above to implement the approach laid out in Section 2. I begin with data on voluntary job changers to test the hypothesis that $F(\cdot)$ is exponential. I then examine whether the wage losses of involuntary job changers are consistent with this estimate of $F(\cdot)$. I close with a brief discussion of the results.

4.1. Is the Offer Distribution Exponential?

Recall from Section 2 that a distinguishing feature of the exponential distribution is that the average wage gain for a voluntary job changer is independent of n. To determine if this condition is satisfied in the data, define for each integer $n \geq 1$ the variable $D_t^{n,n+1}$ which equals 1 if the worker was observed to be on his n-th job at date $t-1$ and on his $n+1$-th job at date t. In addition, define $D_t = \sum_{n \geq 1} D_t^{n,n+1}$ which equals 1 if the worker changed jobs voluntarily between dates $t-1$ and t.

Recall from Section 2 that when $F(\cdot)$ is exponential, we can proceed as if $\gamma = 0$. Thus, we can set γ in (2.1) to zero. In addition, I allow Φ_{it} to contain a quadratic term in potential experience to capture concavity in the wage-experience profile. This implies that the wage growth of job stayers and voluntary job changers can be combined into the following single equation

$$\Delta \ln w_{it} = \beta \Delta X_{it} + \beta_2 \Delta X_{it}^2 + \delta D_t + \sum_{n \geq 2} \delta_n D_t^{n,n+1} + \eta_{it} \qquad (4.1)$$

where

$$\begin{aligned}
\delta &= E\left(R_2 - R_1 \mid N > 1\right) \\
\delta_n &= E\left(R_{n+1} - R_n \mid N > n\right) - \delta \\
\eta_{it} &= \sum\nolimits_{n \geq 1} \left[\left(R_{n+1} - R_n\right) - \left(\delta + \delta_n\right)\right] D_t^{n,n+1} + \Delta\varepsilon_{it}
\end{aligned}$$

The coefficient δ is equal to the average increase in match quality for workers who voluntarily quit their first job, and $\delta + \delta_n$ is equal to the average increase in match quality for workers who quit their n-th job for $n \geq 2$. If $F(\cdot)$ is exponential, then $\delta_n = 0$ for all $n \geq 2$, and δ is the mean of $F(\cdot)$. Conversely, if γ truly is 0, the analysis in Section 2 implies that if $\delta_n = 0$ for all $n \geq 2$, then, up to a location shift, $F(\cdot)$ is exponential with mean δ.

Combining job stayers and voluntary job changers allows us to obtain unbiased estimates for β, β_2, δ, and δ_n using ordinarily least squares (OLS). While the error term η_{it} in (4.1) has zero mean, it is not homoskedastic, since the variance of η_{it} will be greater for job changers than for job stayers. In the special case where $F(\cdot)$ is exponential and $\Delta\varepsilon_{it}$ are i.i.d., the heteroskedasticity takes on a particularly simple form,

$$\sigma_\eta^2 = \begin{cases} \sigma_{\Delta\varepsilon}^2 & \text{if } D_t = 0 \text{ (job stayers)} \\ \sigma_{\Delta\varepsilon}^2 + \lambda^2 & \text{if } D_t = 1 \text{ (job changers)} \end{cases} \tag{4.2}$$

where σ_x^2 denotes the variance of variable x and λ denotes the mean of the exponential distribution. For alternative distributions $F(\cdot)$, the difference $R_{n+1} - R_n$ (conditional on $N > n$) will not be identically distributed across observations, and the heteroskedasticity of η_{it} can be quite arbitrary. For this reason, I account for heteroskedasticity by computing robust (Huber) standard errors that do not require knowing the underlying structure for heteroskedasticity. Robust standard errors have the advantage of also addressing any potential serial correlation in $\Delta\varepsilon_{it}$. I did experiment with estimating (4.1) using feasible generalized least squares (FGLS) under (4.2), and it yielded nearly identical estimates.[12] The formulation of (4.1) makes it easy to check whether the wage gain of those who voluntarily left their n-th job differs significantly from the gain of those who left their first job, i.e. whether $\delta_n = 0$ for a given n. Of course, the true test for whether $F(\cdot)$ is exponential requires simultaneously testing whether $\delta_n = 0$ for all $n \geq 2$.

[12] To be precise, I first estimated (4.1) using OLS and then used the residuals from this regression to estimate $\sigma_{\Delta\varepsilon}^2$ and $\sigma_{\Delta\theta}^2$. I then re-estimated (4.1) by weighted least squares using these variances. In principle, $\sigma_{\Delta\theta}$ should equal λ, the average wage gain for a voluntary job changer. In practice, it is equal to 0.13, while the average gain from voluntary job search is 0.07, although the former is not very precisely measured. One potential explanation for the large point estimate of $\sigma_{\Delta\theta}$ is heterogeneity in λ across workers, which is discussed below.

The results are reported in Table 2. Since there are few observations for workers with very high values of n, I confine my analysis to $n \leq 7$. The number of workers who are observed to change from their n-th to their $n+1$-th job is reported next to the corresponding dummy variable. The first column of coefficients in Table 2 reports estimates for β and β_2 based only on job stayers. The second column combines job changers and job stayers to report estimates for β, β_2, δ, and δ_n. Note that the estimates for β and β_2 are sensitive to combining stayers and voluntary changers, although neither change is statistically significant. The point estimate for δ suggests that the average worker who moves from his first job to his second job gains about 7% on average. The estimated coefficients δ_n for $2 \leq n \leq 7$ are all indistinguishable from zero at the 5% level, i.e. we cannot reject the hypothesis that a worker gains 7.0% on average when he moves from his n-th job to his $n+1$-th job for $2 \leq n \leq 7$. In part, these low t-statistics are due to the small sample sizes for high values n. But it is noteworthy that when sample sizes are greatest, the δ_n are estimated to be near zero with a relatively high degree of precision. To test the joint hypotheses that $\delta_n = 0$ for all $n \geq 2$, I use a robust Wald test that allows for heteroskedastic errors. As reported in the last column of Table 2, the probability that the F-statistic is greater than or equal to the value I calculate is over 70% under the null hypothesis. Thus, I fail to reject the null that $\delta_n = 0$ for all n in the sample. The average wage gain per voluntary job change of 7.4% is on par with Topel and Ward's estimate of an average wage gain of 9.4% per job change among young workers.[13]

Since the failure to reject the hypothesis that $F(\cdot)$ is exponential could be due to the weak power of the test, I consider additional tests to check the hypothesis that $F(\cdot)$ is exponential. A key feature of the exponential distribution is that $(R_{n+1} - R_n)$ is independent of R_n. In fact, as demonstrated in Pfeifer (1982), this feature uniquely identifies the exponential distribution even in richer settings in which observations are no longer identically distributed. Thus, if $F(\cdot)$ is exponential, any variable Z_{it} that is correlated with R_n should not help to predict the wage gain of the worker. This implies that if we estimate

$$\Delta \ln w_{it} = \beta \Delta X_{it} + \beta_2 \Delta X_{it}^2 + (\delta + \delta_Z Z_{it}) D_t + \eta_{it} \tag{4.3}$$

we should observe $\delta_Z = 0$ if $F(\cdot)$ were truly exponential.

Table 3 reports the results for several candidate Z_{it} that are likely to be correlated with R_n. In

[13]Since Topel and Ward combine voluntary and involuntary job changes, it is surprising that their estimate is higher than the one I obtain for only voluntary job changers. This concern is mitigated, although only in part, by the fact that the majority of recorded job changes appear to be voluntary (roughly two thirds of all job changes are reported as voluntary in the NLSY), and that a large share of workers who change jobs involuntarily leave a job where $n = 1$ as opposed to a higher value of n for which the loss would be greater.

all cases, I use deviations from the sample mean of the variable Z. The first column of Table 3 considers n itself. This is equivalent to the test reported in Table 2, except that the effect of n is restricted to be linear. The coefficient on n is clearly negligible, echoing the results in Table 2. The second column examines whether the wage gains of voluntary job changers vary systematically with age. In the absence of involuntary job changers, age would be positively correlated with match quality. While this is mitigated by involuntary job changes (when recall is not possible), it might still be correlated enough with R_n to make for an informative test. As can be seen from the table, age is statistically insignificant in predicting the wage gains for voluntary job changers. I experimented with using a spline in age to see whether there might be some non-linear relationship between age and the average wage gains of voluntary job changers, but this specification also proved statistically insignificant.

Column (3) considers potential experience, which is arguably a better proxy for R_n than age, although it too might be only weakly correlated with R_n given workers experience involuntary job changes. The coefficient on potential experience is negative, in line with the findings in Topel and Ward (1992) that the average wage gain per job change declines with potential experience. But this coefficient is statistically insignificant at conventional levels. Since potential experience is a linear combination of age and education, I also allow the two terms to enter separately to see whether these results might be masking or reflecting differences across education groups. The results in column (4) confirm the inability of age to predict the wage gains of voluntary job changers, but suggest that wage gains do vary significantly by education.

The observation that wage gains vary by education groups is incompatible with my assumption that all workers face the same essential search problem. One possibility is that workers face the same offer distribution $F(\cdot)$ but that workers with different educational backgrounds face different search probabilities p and s. In this case, the distribution of n will differ across education groups, so that the average match quality of a random job changer with a higher educational background will be different from one with a lower educational background. This would provide evidence that $F(\cdot)$ is not exponential, since it would imply that workers gain different amounts on average depending on the current value of R_n. But if this were really the reason for why wage gains vary systematically across education groups, we should have also observed differences in average wage gains by n, contrary to the results in Table 2.

A more likely explanation is that workers from different education groups sample from different distributions $F(\cdot)$. In principle, we could still carry out the analysis described in Section 2 for each education group separately. However, the sample sizes in the NLSY make this impractical. This raises the question of how much we can learn from aggregate data. In particular, can we conclude from the fact that the average wage gain across all workers does not vary with n that $F(\cdot)$ is likely to be exponential for each education group, albeit with different means? While it is true that if $F(\cdot)$ is exponential for each group then the average wage gain *within* each education group would be constant for all n, it might not be true that the average wage gain *across* education groups is constant for all n. In particular, it will not be true if the distribution of educational attainment among voluntary job changers varies with n. To assess the importance of these composition effects, Figure 2 illustrates the distribution of educational attainment in my sample for workers who move from their n-th to their $n+1$-th job for various n. There are somewhat more workers with low educational attainment among those who quit their first job than among those who quit their second job. But this difference is small, and the composition across voluntary job changers are nearly identical for $n \geq 2$. This suggests composition bias should play only a minor role, so that the fact that the average wage gain across all workers is constant across n is a necessary condition for each $F(\cdot)$ to be exponential. However, the converse is not true: the average wage gain *within* each group might differ by n, so that each group must face an offer distribution that is not exponential, but these differences could still potentially cancel out *across* education groups. Such a coincidence seems unlikely, but there is nothing that inherently rules it out.

Finally, in the fifth column of Table 3, I examine whether the wage gains for voluntary job changers vary with the completed tenure on the job the worker voluntarily left. This tenure variable is likely to be correlated with the quality of the match θ_{ij} on the job the worker left, since workers will tend to stay longer in matches of higher quality.[14] The results imply that we can reject the null hypothesis that tenure on the job the worker left is uncorrelated with the wage loss at the 1% level. While this result is troubling, there is an alternative interpretation of the negative correlation that is consistent with the underlying offer distribution being exponential. Recall from Jovanovic (1979) and Flinn (1986) that if workers and employers only get to observe θ_{ij} with noise, then workers will be more likely to leave their job the longer they stayed with that job, other things equal. This is because the option value of remaining in a match declines the more certain they are about the quality of their match. Thus, workers who spent more time with their previous employer would be less selective in accepting new offers, and their wage gains would appear smaller on average.

[14] However, as noted by Topel (1991), this is not necessarily true once we condition on potential experience. This is because within a set amount of time, workers who voluntarily changed jobs more frequently are likely to have both low tenure and high match quality. Since the NLSY follows a single cohort with roughly similar experience, this caveat is certainly warranted. But my test only requires that my measure be correlated with match quality, not that it be positively correlated.

In sum, the average wage gains of voluntary job changers do not appear to vary significantly with measures that are likely to be correlated with previous match quality, such as the number of jobs n the worker previously worked on or the time since the worker entered the labor market. This observation is true for the exponential distribution, and in some circumstances uniquely characterizes this distribution. However, the small sample sizes raise questions about the power of these tests, and the fact that wage gains are related to the tenure on the worker's previous job is inconsistent with the assumption that $F(\cdot)$ is exponential in the basic version of the model. But as noted in Section 2, the wage losses of involuntary job changers offer an additional over-identifying restriction that can be used to check if $F(\cdot)$ is exponential. I now turn to this prediction.

4.2. Accounting for the Wage Losses of Displaced Workers

Given an offer distribution $F(\cdot)$ and an estimate the returns to tenure γ, we can directly compute the average wage loss of a worker displaced from his n-th job as implied by the model. Comparing this to the average losses in the data should therefore offer a consistency check on our estimate of $F(\cdot)$. I begin with the task of estimating the returns to seniority γ, and then turn to the task of computing implied average wage loss given the distribution $F(\cdot)$ estimated above.

As noted above, previous authors have already tackled the question of estimating returns to seniority, although the results are still in dispute. Early work by Altonji and Shakotko (1987) and Abraham and Farber (1987) found fairly modest returns to tenure. Subsequent work by Topel (1991) challenged these results and proposed an alternative methodology which reveals much more substantial returns to tenure. Altonji and Williams (1997) have recently attempted to reconcile the two approaches, and argue that estimates lie somewhere inbetween those reported in the two papers, albeit closer to the original smaller estimates. All of the papers above use data from the Panel Survey of Income Dynamics (PSID), which involves older and more established workers than the ones I consider. For example, the average potential experience in my sample is 9.8 years, compared with about 20 years in the papers cited above. This distinction is important, since young workers are more mobile and thus have less incentive to invest in skills specific to any one particular job. For this reason, we would want to directly estimate average returns to tenure in the NLSY rather than rely on any of these previous estimates.

One exception that does look at NLSY data is Altonji and Williams (1998). They develop an alternative Bayesian approach to estimate returns to tenure, and include the NLSY dataset as part of their analysis. They find statistically significant but economically small returns to tenure. As such, assuming $\gamma = 0$ would appear to provide a reasonable first approximation to the true γ. To confirm this, I now turn to the methodology advanced in Topel (1991), which seems to yield the largest estimates for returns to tenure. As I argue next, even this approach yields modest returns to tenure in my sample.

Topel's approach uses the fact that we can rewrite $X_{it} = X_{0it} + T_{it}$, where X_{0it} denotes the experience of the worker when he first started working on job $j(t)$. Using this, we can rewrite the wage equation as

$$\ln w_{it} = \beta X_{0it} + (\beta + \gamma) T_{it} + \beta_2 X_{it}^2 + \phi_i + \theta_{it} + \varepsilon_{it} \tag{4.4}$$

But as can be seen in (2.1), we can easily obtain an unbiased estimate of $\beta + \gamma$ from the wage growth of job stayers. Thus, Topel suggests a two stage regression, where in the first step we estimate $\beta + \gamma$ and β_2 from the wage growth of job stayers, and in the second stage we regress $\ln w_{it} - (\beta + \gamma) T_{it} - \beta_2 X_{it}^2$ on X_{0it}. Topel reasons that X_{0it} is likely to be positively correlated with θ_{it}, since workers are more likely to arrive at good matches over time. Hence, the second stage estimate for β should be biased upwards. On the other hand, Altonji and Williams (1997) argue that X_{0it} is likely to be negatively correlated with ϕ_i, implying the estimate for β should be biased downwards.[15] To account for the latter, I allow for individual fixed effects in the second-stage regression. As to the former, as noted in Altonji and Shakotko (1987), involuntary job loss will weaken the correlation between potential experience and match quality, suggesting the bias is likely to be small. This is affirmed by evidence in Figure 1 that the distribution of n settles down to an invariant distribution, which suggests that time spent in the labor market is likely to be only weakly correlated with search capital in my sample.

Table 4 reports the results of the two-step approach. The first entry reports the estimate of $\beta + \gamma$, which is just the coefficient on ΔX_{it} from column (1) of Table 2. The point estimate is given 0.0794. By comparison, Topel (1991) estimates this at 0.1258. Since Topel estimates β at 0.0713, the only way for my sample to accord with his estimates of the returns to experience is if γ in my sample is small. This observation is important, since as Altonji and Williams (1997) note, even though "recent literature has been quite divided on the value of seniority, almost all research suggests that the return to general labor market experience is large" (p234). Indeed, as reported in the second entry of Table 4, my estimate for β in the second stage regression is given

[15] Topel (1991) also mentions this possibility, but argues its role is small in his sample. But this is more likely to be a problem in the NLSY dataset, which follows a single cohort over time. In particular, within a fixed cohort, better workers are more likely to obtain more schooling and thus have less potential experience.

by 0.0740, which is similar to Topel's estimate. The implied point estimate for γ is 0.0054, i.e. half a percentage point per year. If anything, this estimate is even smaller than reported in Altonji and Williams (1998). Still, the standard error of this estimate (which is corrected to reflect sampling error in the first stage regression) implies γ is significantly different from zero at the 5% level. But the magnitude is clearly small enough that we can probably abstract from these returns in estimating the wage loss of involuntary job changers.

The second row in Table 4 reports my estimated returns to tenure and potential experience at different horizons. Returns to tenure are by assumption linear and thus grow at a constant rate over time. The estimated returns to experience appear concave, and their magnitudes are consistent with those reported in Altonji and Williams (1998). For example, the cumulative returns to experience according to my estimates are 0.327 and 0.568 log points after 5 and 10 years, respectively, compared to 0.321 and 0.560 in Altonji and Williams (1998). Topel's (1991) estimates for the cumulative returns to experience in the PSID are somewhat smaller, at 0.268 and 0.414, respectively. But at low levels of experience, where my estimates are likely to be more precise given the young age of workers in my sample, the two estimates are quite close.

Table 4 therefore suggests that assuming $\gamma = 0$ might be reasonable for my sample, so that the bound I computed in Section 2 under this assumption provides a good approximation of the true lower bound. Before turning to this bound, though, notice that I can use Topel's approach to test an additional assumption implicit in my model, namely that returns to tenure are linear. In particular, we can add a $\gamma_2 \Delta T_{it}^2$ term to the first stage regression and check whether γ_2 is significantly different from zero. This is reported in the last line of Table 4. The coefficient γ_2 does prove to be significantly different from zero at the 1% level, suggesting returns to tenure do accumulate at a non-linear rate, in violation of my assumptions. But again, the magnitudes involved are relatively small, especially since tenure on most jobs in my sample is short, in which case no returns to seniority should still provide a reasonable approximation.

Recall from Section 2 that when $\gamma = 0$, we can use bounds on conditional record moments from an exponential parent distribution $F(\cdot)$ with mean λ to generate the following restriction on the average wage loss for workers who lose their n-th job:

$$E\left(\left|\Delta \ln w_{it} - \beta \Delta X_{it} - \beta_2 \Delta X_{it}^2\right| \;\middle|\; N_{t-1} = n\right) \geq \lambda \left(\sum_{i=1}^{n} \frac{1}{i} - 1 \right) \equiv \lambda \left(g\left(n\right) - 1\right) \qquad (4.5)$$

Once again, then, let $D_t^{n,1}$ equal 1 if the worker was on his n-th job at date $t-1$ but was on a job where $n = 1$ at date t. In the case where $\gamma = 0$, we can rewrite (2.1) as

$$\Delta \ln w_{it} = \beta \Delta X_{it} + \beta_2 \Delta X_{it}^2 - \sum\nolimits_{n \geq 1} \delta_n D_t^{n,1} + \eta_{it} \tag{4.6}$$

where

$$\begin{aligned} \delta_n &= E(R_n \mid N_{t-1} = n) - E(R_1) \\ \eta_{it} &= \sum_{n \geq 1} [(R_n - R_1) - \delta_n] D_t^{n,1} + \Delta \varepsilon_{it} \end{aligned}$$

We can thus estimate δ_n using OLS and then confirm whether $\delta_n \geq \lambda(g(n) - 1)$ as implied by the model.

Figure 3 displays the point estimates for each δ_n graphically, together with a 95% confidence interval for each estimate and the value of the lower bound $\lambda(g(n) - 1)$ for $\lambda = 0.0737$. Although my standard error bands are quite wide given the small sample sizes involved, for two values of n we can reject the null hypothesis that workers do not experience any wage losses at the 5% and 10% significance level, respectively. This provides some evidence that involuntary job loss involves some loss of search capital. But the most striking aspect of Figure 3 is that actual average wage losses are quite close to the lower bound. This suggests there is little evidence to suggest that wage losses fail to satisfy the lower bound imposed by the model. Moreover, to the extent that observed wage losses exceed this bound, it is not by a wide margin. This last observation confirms that returns to tenure cannot be very large, since the wage losses of involuntary job changers net of returns to experience β that accord with most previous estimates are never much more than what is already implied by forgone search capital.

To formally test the bound restriction in (4.5), it will prove convenient to renormalize the dummy variables $D^{n,1}$ by multiplying them by $(g(n) - 1)$ for each $n \geq 2$. That is, we can rewrite (4.6) as

$$\Delta \ln w_{it} = \beta \Delta X_{it} + \beta_2 \Delta X_{it}^2 - \widetilde{\delta}_1 D_t^{1,1} - \sum\nolimits_{n \geq 2} \widetilde{\delta}_n \left[(g(n) - 1) \cdot D_t^{n,1} \right] + \eta_{it} \tag{4.7}$$

where

$$\widetilde{\delta}_n = \begin{cases} \delta_n & \text{if } n = 1 \\ \delta_n (g(n) - 1)^{-1} & \text{if } n \geq 2 \end{cases}$$

It follows that

$$\delta_n \geq \lambda(g(n) - 1) \Leftrightarrow \begin{cases} \widetilde{\delta}_1 \geq 0 & \text{if } n = 1 \\ \widetilde{\delta}_n \geq \lambda & \text{if } n \geq 2 \end{cases} \tag{4.8}$$

Note the difference between workers who are separated from their first job and workers who are separated from a job where $n \geq 2$. The lower bound for workers who lose their first job is independent of the underlying distribution. Instead, the bound is equal to $E(\theta_{1:1}) - E(\theta_{1:1})$, which is equal to 0 for any $F(\cdot)$. On the one hand, this implies that workers who are displaced from their first job afford a way to test an implication of the model that is independent of the underlying distribution of match quality, albeit a weak one. At the same time, it also implies that the wage losses of these workers cannot be used to identify λ, and thus are not useful for distinguishing whether the underlying distribution $F(\cdot)$ is exponential.

The first three columns of Table 5 report results for testing the hypothesis given by (4.8). The coefficients are estimated using OLS, with robust standard errors to adjust for heteroskedasticity. For $n = 1$, the point estimate for $\widetilde{\delta}_1$ is close to its theoretical bound of zero, and we cannot reject the null that $\widetilde{\delta}_1 \geq 0$. For $n = 2$, 4, and 7, the point estimate of $\widetilde{\delta}_n$ is actually quite close to the theoretical lower bound of $\lambda = 0.0737$. Note that for all $2 \leq n \leq 7$, we fail to reject the null hypothesis in (4.8) at the 5% level (using a one-sided test), even in the most extreme cases where the point estimate for $\widetilde{\delta}_n$ is negative. A joint Wald test of (4.8) for all $n \leq 7$ also passes at a comfortable level of significance. Thus, the wage gains associated with voluntary job changes line up with the wage losses observed among involuntary job changes, in the sense that both are consistent with $F(\cdot)$ being exponential with mean 0.0737.

If we further maintain the hypothesis $F(\cdot)$ is exponential, we can get a more efficient test of whether the average wage losses of involuntary job changers are consistent with the average wage gains of voluntary job changers. In particular, if $F(\cdot)$ is exponential, regressing the wage losses of involuntary job changers on $(g(n) - 1)$ should yield an upwardly biased estimate of λ, while the wage gains of voluntary job changers should yield an unbiased estimate for λ. Moreover, given the relatively high rate of involuntary job turnover among young workers, which would imply low values for α, we would expect the bias in the first regression to be small, so that both approaches should yield very similar estimates for λ. Columns 4-9 of Table 5 report the estimate for the coefficient on $\sum (g(n) - 1) D_{n,1}$ over different values of n. In general, my estimates are actually lower than 0.0737, the estimate I get for voluntary job changers. However, if I limit attention to $n \leq 2$, where sample sizes are largest, the estimate for λ among involuntary job changers exceeds 0.0737. But even when I include all observations, the coefficient for λ is not statistically different from this value. Thus, the wage losses we observe among involuntary job changers are consistent with the exponential distribution that is suggested in the previous section, although admittedly sample sizes in my dataset are too small to definitively assess the model's fit.

4.3. Discussion

To recap, the analysis above suggests that a search model with no recall in which workers draw log wage offers from a fixed exponential distribution can reasonably account for both the wage growth of workers who change jobs voluntarily and the wage loss of workers who change jobs involuntarily. The fact that a search model can replicate the dynamic path of wages complements recent work that argues search models can successfully replicate other dimensions of wage data, namely the cross-section of wages. Moreover, the fact that my results suggest a Pareto wage offer distribution accords with previous work that suggest this distribution has a Pareto tail.

At the same time, my results appear to conflict with certain established results in the literature. For example, McCall (1990) and Neal (1999) argue, using the same NLSY data, that young workers tend to search in stages, first settling on an occupation and then searching for an employer within an occupation. It might therefore seem surprising that displaced workers have to start from scratch each time they lose their job, since they already learned about which occupation they are most suited for and would only search in that occupation once displaced. But upon further inspection, my results might not be as inconsistent with these findings as might seem at first. This is because workers tend to settle on an occupation fairly quickly; for example, Neal reports that 70% of workers who leave their first job in a given career change to a different career, but only 20% of those who leave their second job in a given career change to a different career. If the majority of workers in my sample are already searching for a job within a given career, it is not surprising that they would appear to resume searching from scratch when they are displaced. This is particularly relevant given the way I construct my sample; since I only assign n if the worker experienced some turnover, specifically at least one involuntary job change, the workers in my sample are more likely to have already learned something about which career they are most suited for.

The fact that workers who are displaced from their first job (where $n = 1$) appear to suffer very little, i.e. $\delta_1 \approx 0$, is interesting in its contrast to evidence of large and persistent losses documented for displaced workers by Ruhm (1991) and Jacobson, Sullivan, and LaLonde (1993), among others. Of course, these papers focus on established workers, who are both more likely to be on a high quality match and might have invested a fair amount in match-specific human capital, so there is nothing contradictory about the fact that workers with $n = 1$ suffer very little from involuntary job loss. For workers with $n > 1$, by contrast, the model does imply involuntary job loss will involve some wage loss on average, a point that is largely confirmed by the point estimates. Given an estimate of the offer distribution $F(\cdot)$, we can directly compute these losses, as well as how long it takes a worker to rebuild forgone search capital, to get a sense of how important search capital might be in accounting for the wage losses of more established workers. For the exponential distribution I estimate above, the average wage loss among workers with $n \geq 4$ (who make up about 20% of the sample) is at least 10%, in line with estimates in some of the papers cited above. As for the persistence of these losses, note that the number of offers it takes a worker until he finds a match of quality θ has a geometric distribution with a success probability of $1 - F(\theta)$. For the exponential case, it follows that $(1 - F(\theta))^{-1} = \exp(\theta/\lambda)$. Since θ is the log of the wage paid on a given job, it follows that the average number of offers is linear in the wage *level*. In other words, on average it takes twice as many offers for the worker to get a job that pays a wage that is twice as large. Thus, rebuilding search capital may not necessarily be particularly quick given the offer distribution is so heavily concentrated towards the low end of the support of the distribution.

5. Conclusion

This paper attempted to gauge the empirical relevance of models in which workers search for high paying jobs by repeatedly drawing offers from a fixed offer distribution. In particular, it shows how one can use wage data to identify the offer distribution if the data were in fact generated by such a model. The results suggest that the same distribution we identify from voluntary job changers is consistent with the wage losses we observe for involuntary job changers given their previous work history. In fact, at least among the young workers in my sample, it appears as if search considerations can account for almost all of the wage growth among young workers beyond that which can be attributed to general human capital. This finding seems reasonable; young workers have less incentive to invest in match-specific human capital than older and more established workers who anticipate lower future mobility, and there is already some evidence, albeit contested, which suggests returns to tenure are small even among established workers.

While the particular application emphasized in this paper concerns quantifying the role of search capital, once we identify the underlying offer distribution $F(\cdot)$, we can potentially use it to carry out various simulations and policy experiments. For example, in Barlevy (2002), I use the offer distribution identified in this paper to estimate the effects of macroeconomic shocks on the allocation of workers and their implications for aggregate productivity.

More generally, the analysis above demonstrates the natural connection between search models and record statistics. My particular analysis exploits characterization theorems that can relate record moments back to a unique parent distribution. But record-value theory contains a wealth of results that could similarly be used in analyzing search models. For example, results on the frequency of records in models with and without drift could be used to test for the presence of tenure effects directly from observations on job mobility, without any reference to wage data. This relies on the idea that records, and hence turnover, should be more infrequent if workers accumulate match-specific human capital. This test is analogous to the way statisticians have suggested using the frequency of record temperatures to detect shifts in the distribution of temperatures that are indicative of global warming. As for wage data, this paper has largely focused on the case in which there are no returns to seniority, i.e. where $\gamma = 0$. In future work, I hope to tackle the more interesting but also more challenging case of positive returns to tenure, including particular cases of non-linear returns that are simple enough to analyze tractably.

Beyond job search models, results from record-value theory are likely to be useful for other economic applications. For example, as Kortum (1997) already notes, search models can be similarly applied to studying research and development and patenting activities. Beyond applications that directly involve search, statisticians have used record models to study optimal stopping rules, as discussed in some detail in Arnold, Balakrishnan, and Nagaraja (1998). Other non-search applications include insurance with imperfect commitment, e.g. Beaudry and DiNardo (1991), where the optimal contract implicitly conditions transfers on the occurrence of record-setting events. Thus, adapting and extending results in record-value theory can provide economists with useful tools with which to analyze a variety of dynamic models.

6. Appendix

Proof of Remark 2: By conditioning on the number of periods before separation, we have

$$\begin{aligned}
\Pr(M=m) &= \sum_{j=m}^{\infty}\binom{j}{j-m+1}(1-p)^{j-m+1}p^{m-1}(1-s)^j s \\
&= \frac{d^{m-1}}{d(1-p)^{m-1}}\left[\sum_{j=1}^{\infty}(1-p)^j(1-s)^j\right]\frac{sp^{m-1}}{(m-1)!} \\
&= \frac{d^{m-1}}{d(1-p)^{m-1}}\left[\frac{(1-p)(1-s)}{p+(1-p)s}\right]\frac{sp^{m-1}}{(m-1)!} \\
&= \frac{(m-1)!(1-s)^{m-1}}{(m-1)!(p+(1-p)s)^m}sp^{m-1} \\
&= \left(\frac{(1-s)p}{p+s-ps}\right)^{m-1}\frac{s}{p+s-ps} \\
&= (1-\alpha)^{m-1}\alpha
\end{aligned}$$

which completes the proof. ■

Proof of Lemma: According to Bunge and Nagaraja (1991), the probability density function of the first $n+1$ records in the geometric random records model is given by

$$h(r_1, r_2, ..., r_{n+1}; N > n) = f(r_{n+1})\prod_{i=1}^{n}\frac{qf(r_i)}{1-qF(r_i)} \tag{6.1}$$

where $f(\cdot)$ denotes the pdf of θ_m, $F(\cdot)$ denotes the cdf of θ_m, and $q = 1-\alpha$. Next, note that for any cdf $F(\cdot)$, the random variable $F(R_n)$ has the same distribution as U_n, where U_n is the n-th record in a geometric random record model where θ_m is distributed uniformly over $[0,1]$. Using the fact that for the uniform distribution, $f(u) = 1$ and $F(u) = u$, we can get an expression for $h(u_{n+1}; N > n)$ by integrating out u_1 through u_n in (6.1). Using a standard induction argument, we can establish that

$$h(u_{n+1}; N > n) = \frac{[-\ln(1-qu_{n+1})]^n}{n!}$$

Define the inverse cdf $F^{-1}(x)$ for $x \in (0,1)$ as $\sup\{y : F(y) \le x\}$. Then the expected value of $|R_{n+1}|$ conditional on $N > n$ is given by

$$\begin{aligned}
E\left(\ |R_{n+1}|\ \middle|\ N > n\right) &= \int_0^1 \left|F^{-1}(u)\right|\frac{[-\ln(1-qu)]^n}{n!\Pr(N>n)}du \\
&\le \frac{[-\ln(1-q)]^n}{n!\Pr(N>n)}\int_0^1\left|F^{-1}(u)\right|du \\
&= \frac{[-\ln(1-q)]^n}{n!}E(|\theta_m|) < \infty
\end{aligned}$$

Since $E\left(\ |R_n|\ \middle|\ N>n\right) < E\left(\ |R_{n+1}|\ \middle|\ N>n\right)$, the former is also finite. Since $E(|X|)<\infty$ implies $E(X)<\infty$, the lemma follows. ■

Proof of Proposition: Integrating out (6.1) yields the following densities:

$$\begin{aligned} h(r_{n+1}, r_n; N>n) &= f(r_{n+1}) \frac{[-\ln(1-qF(r_n))]^{n-1}}{(n-1)!} \frac{qf(r_n)}{1-qF(r_n)} \\ h(r_n; N>n) &= [1-F(r_n)] \frac{[-\ln(1-qF(r_n))]^{n-1}}{(n-1)!} \frac{qf(r_n)}{1-qF(r_n)} \end{aligned}$$

Next, define $\Delta = r_{n+1} - r_n$. By construction, $\Delta \geq 0$. Using the law of iterated expectations,

$$\begin{aligned} E(\Delta \mid N \geq n) &= E(E(\Delta \mid r_n, N>n)) \\ &= E\left(\int_0^\infty \Delta\ h(\Delta \mid r_n, N \geq n)\, d\Delta\right) \end{aligned}$$

where $h(\Delta \mid r_n, N \geq n)$ is the density of the difference between the n-th record and the $n+1$-th record conditional on r_n and that there are more than n records. Note that

$$\begin{aligned} h(\Delta \mid r_n, N \geq n) &= \frac{h(r_n+\Delta, r_n \mid N>n)}{h(r_n \mid N>n)} \\ &= \frac{f(r_n+\Delta)}{1-F(r_n)} \end{aligned}$$

So that the conditional expectation of Δ is given by

$$\begin{aligned} E(\Delta \mid r_n, N \geq n) &= \int_0^\infty \Delta \frac{f(r_n+\Delta)}{1-F(r_n)} d\Delta \\ &= \frac{\int_0^\infty [1-F(r_n+\Delta)]\, d\Delta}{1-F(r_n)} \end{aligned}$$

and if we integrate the above expression over r_n, we have

$$\begin{aligned} E(\Delta \mid N>n) &= E\left(\frac{\int_0^\infty [1-F(r_n+\Delta)]\, d\Delta}{1-F(r_n)} \middle|\ r_n, N>n\right) \\ &= \int_{-\infty}^\infty \left[\frac{\int_0^\infty [1-F(r_n+\Delta)]\, d\Delta}{1-F(r_n)}\right] \frac{h(r_n; N>n)}{\Pr(N>n)} dr_n \\ &= \int_{-\infty}^\infty \left[\frac{\int_0^\infty [1-F(r_n+\Delta)]\, d\Delta}{1-F(r_n)}\right] \frac{[1-F(r_n)]\, qf(r_n)}{1-qF(r_n)} \frac{[-\ln(1-qF(r_n))]^{n-1}}{(n-1)!} dr_n \\ &= \int_{-\infty}^\infty \left[\int_0^\infty [1-F(r_n+\Delta)]\, d\Delta\right] \frac{qf(r_n)}{1-qF(r_n)} \frac{[-\ln(1-qF(r_n))]^{n-1}}{(n-1)!} dr_n \end{aligned}$$

Now, suppose we have two functions F_1 and F_2 such that

$$E\left(R_{n+1}^{(1)} - R_n^{(1)} \mid N > n\right) = E\left(R_{n+1}^{(2)} - R_n^{(2)} \mid N > n\right)$$

for $n = 1, 2, 3, ...$ Then we have

$$\int_{-\infty}^{\infty} \left[\int_0^{\infty} \left[1 - F\left(r_n + \Delta\right)\right] d\Delta\right] \frac{\left(-\ln\left(1 - qF_1\left(r_n\right)\right)\right)^{n-1}}{(n-1)!\left(1 - F_1\left(r_n\right)\right)} \frac{qf_1\left(r_n\right)}{1 - qF_1\left(r_n\right)} dr_n =$$
$$\int_{-\infty}^{\infty} \left[\int_0^{\infty} \left[1 - F\left(r_n + \Delta\right)\right] d\Delta\right] \frac{\left(-\ln\left(1 - qF_2\left(r_n\right)\right)\right)^{n-1}}{(n-1)!\left(1 - F_2\left(r_n\right)\right)} \frac{qf_2\left(r_n\right)}{1 - qF_2\left(r_n\right)} dr_n$$

Rewrite both integrals using the change of variables $u = F\left(r_n\right)$ to get

$$\int_0^1 \left[\int_0^{\infty} \left[1 - F_1\left(F_1^{-1}\left(u\right) + \Delta\right)\right] d\Delta\right] \frac{\left(-\ln\left(1 - qu\right)\right)^{n-1}}{(n-1)!\left(1 - u\right)} \frac{q}{1 - qu} du =$$
$$\int_0^1 \left[\int_0^{\infty} \left[1 - F_2\left(F_2^{-1}\left(u\right) + \Delta\right)\right] d\Delta\right] \frac{\left(-\ln\left(1 - qu\right)\right)^{n-1}}{(n-1)!\left(1 - u\right)} \frac{q}{1 - qu} du$$

Using Lemma 3 in Lin (1987) page 480, we know that if

$$\int_0^1 f\left(x\right)\left(-\ln\left(1 - x\right)\right)^n dx = 0$$

for all $n \geq 1$ then $f\left(x\right) = 0$ for all $x \in (0, 1)$. By a simple contradiction argument, one can show that this implies that if

$$\int_0^1 f\left(x\right)\left(-\ln\left(1 - qx\right)\right)^n dx = 0$$

for all $n \geq 1$, then $f\left(x\right) = 0$. From this, we can deduce that for any u,

$$\int_0^{\infty} \left[1 - F_1\left(F_1^{-1}\left(u\right) + \Delta\right)\right] d\Delta = \int_0^{\infty} \left[1 - F_2\left(F_2^{-1}\left(u\right) + \Delta\right)\right] d\Delta$$

Let $t = F_1^{-1}\left(u\right) + \Delta$. Then it follows that for any u,

$$\left[\int_{F_1^{-1}(u)}^{\infty} \left[1 - F_1\left(t\right)\right] dt\right] = \left[\int_{F_2^{-1}(u)}^{\infty} \left[1 - F_2\left(t\right)\right] dt\right]$$

Since $F_1\left(\cdot\right)$ and $F_2\left(\cdot\right)$ are continuous, nondecreasing, and bounded, it follows that they are both differentiable almost everywhere. This, in turn, implies that $F_1^{-1}\left(u\right)$ and $F_2^{-1}\left(u\right)$ are differentiable for almost every $u \in (0, 1)$. Differentiating with respect to such u yields

$$\left[1 - F_1\left(F_1^{-1}\left(u\right)\right)\right] \frac{d}{du} F_1^{-1}\left(u\right) = \left[1 - F_2\left(F_2^{-1}\left(u\right)\right)\right] \frac{d}{du} F_2^{-1}\left(u\right)$$

Since $F_1\left(F_1^{-1}(u)\right) = F_2\left(F_2^{-1}(u)\right) = u$, it follows that for almost all $u \in (0,1)$,

$$\frac{d}{du} F_1^{-1}(u) = \frac{d}{du} F_2^{-1}(u)$$

Integrating out yields

$$F_1^{-1}(u) = F_2^{-1}(u) + c$$

for some constant c, which establishes the claim. ■

Table 1: Summary Statistics

	all	invol	vol
average education	12.9	12.3	12.8
fraction of sample:			
below HS	0.20	0.30	0.25
HS	0.34	0.34	0.30
some college	0.26	0.23	0.26
college	0.20	0.13	0.19

observations that involve…	all jobs	in sample
voluntary job change	0.25	0.11
involuntary job change	0.16	0.07

	mean	median
age	27.11	27.00
experience	9.79	10.00
tenure	3.47	3.00
wage (1992 dollars)	$11.28	$9.30

# of obs	34,431
# of individuals	5,404
# of jobs	13,229

Statistics are based on the final sample used in Tables 2-5, except for the number of job changes involving a voluntary job change or an involuntary job change. In that case, both the number of all jobs in the original NLSY data as well as in the final sample are reported

Table 2: The Wage Gains of Voluntary Job Changers, by *n*

	sample size	(1)	(2)	(3)
ΔX (constant)	--	0.0794	0.0828	0.0826
		0.0065	0.0063	0.0065
ΔX^2	--	-0.0017	-0.0019	-0.0019
		0.0003	0.0003	0.0003
D	1,886		0.0695	0.0737
			0.0131	0.0075
D^{23}	916		0.0137	--
			0.0227	
D^{34}	543		0.0093	--
			0.0276	
D^{45}	302		-0.0314	--
			0.0293	
D^{56}	152		0.0670	--
			0.0552	
D^{67}	75		-0.0149	--
			0.0950	
D^{78}	55		-0.0161	--
			0.0845	
# obs		28,015	31,944	31,944
stayers			28,015	28,015
changers			3,929	3,929
				F (6,31395) = 0.61
				Prob > F = 0.7229

The dependent variable is the difference in log real wages (in 1992 dollars using the implicit GDP deflator). The independent variables are the difference in potential experience, which is identically 1, the difference in potential experience squared, which is just 2X-1, a set of dummy variables $D^{n,n+1}$ that equal 1 if the worker moved from his n-th job to his n+1-th job, and a dummy variable D that is the sum of all these dummy variables for all $n \geq 1$. The column labeled sample size denotes the number of workers who voluntarily left their n-th job, where number next to D denotes the number of workers who left their first job. The numbers below the coefficients denote robust standard errors. The F-statistic denotes the robust Wald-statistic comparing the unconstrained regression in column (2) to the constrained regression in the column (3) in which the coefficients on $L^{n,n+1}$ are set to zero for all $n \geq 2$.

Table 3: Explaining the Wage Gains of Voluntary Job Changers

	(1)	(2)	(3)	(4)	(5)
ΔX (constant)	0.0826	0.0804	0.0794	0.0780	0.0810
	0.0063	0.0063	0.0063	0.0063	0.0063
ΔX^2	-0.0019	-0.0018	-0.0017	-0.0016	-0.0018
	0.0003	0.0003	0.0003	0.0003	0.0003
D	0.0737	0.0738	0.0738	0.0739	0.0739
	0.0094	0.0094	0.0094	0.0093	0.0094
D*n	0.0009				
	0.0072				
D*age		-0.0024		-0.0034	
		0.0027		0.0027	
D*exp			-0.0032		
			0.0028		
D*educ				0.0120	
				0.0036	
D*tenure$_{t-1}$					-0.0149
					0.0049
# obs	31,944	31,944	31,944	31,944	31,944
stayers	28,015	28,015	28,015	28,015	28,015
changers	3,929	3,929	3,929	3,929	3,929

The dependent variable is the difference in log real wages (in 1992 dollars using the implicit GDP deflator). The independent variables are as in Table 2, with the addition of interaction terms between the dummy variable D that equals one if the worker quit and variables such as the number of jobs *n* the worker has held since his last involuntary job change, age, potential experience, years of schooling, and tenure on the job the worker quit, respectively. The numbers below the coefficients denote robust standard errors.

Table 4: Estimating Returns to Tenureγ

linear returns to tenure

	within-job wage growth $\beta+\gamma$		experience effect β		tenure effect γ
	0.0794		0.0740		0.0054
	0.0065		0.0061		0.0024
	1 year	2 years	5 years	7 years	10 years
returns to tenure	0.0054	0.0108	0.0271	0.0380	0.0542
	0.0024	0.0049	0.0122	0.0171	0.0245
returns to experience	0.0723	0.1411	0.3270	0.4337	0.5680
	0.0058	0.0109	0.0226	0.0274	0.0300

non-linear returns to tenure

within-job wage growth $\beta+\gamma$	experience effect β	tenure effect γ	tenure squared γ_2
0.0826	0.0661	0.0165	-0.00157
0.0065	0.0067	0.0024	0.00048

Returns to tenure are estimated using the two-step method outlined in Topel (1991). The first stage regression corresponds to column (1) in Table 2. The coefficient on ΔX correponds to the estimate of $\beta+\gamma$ reported above, while the coefficient on ΔX^2 is denoted by β_2 and is not reported here. The second stage involves regressing the log of the real wage net of $(\beta+\gamma)T + \beta_2 X^2$ on initial experience and individual fixed-effects. The coefficient on initial experience corresponds to the estimate of β above, and the difference corresponds to the estimate of γ above. The standard errors are adjusted to reflect estimation error in the first-stage regressor. In particular, they were computed using the stacking and weighting procedure suggested in Altonji and Williams (1998). Returns to tenure and returns to experience are constructed using the estimated coefficients γ, β, and β_2 for different time horizons. The final panel adds the change in tenure squared as an explanatory variable in the first stage regression, whose coefficient corresponds to γ_2, and this term is also subtracted from the real wage in the second stage regression.

Table 5: The Wage Losses of Involuntary Job Changers, by *n*

	(1)	(2) sample size	(3) $Pr(\delta_n \geq \lambda)$	(4) $n \leq 2$	(5) $n \leq 3$	(6) $n \leq 4$	(7) $n \leq 5$	(8) $n \leq 6$	(9) $n \leq 7$
ΔX (const)	0.0829	--	--	0.0811	0.0807	0.0823	0.0818	0.0815	0.0816
	0.0062			0.0062	0.0062	0.0062	0.0062	0.0062	0.0062
ΔX^2	-0.0019	--	--	-0.0018	-0.0018	-0.0019	-0.0019	-0.0019	-0.0019
	0.0003			0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
D*(1-g(n))	--	--	--	0.0998	0.0569	0.0656	0.0468	0.0394	0.0437
				0.0437	0.0311	0.0271	0.0274	0.0276	0.0261
D^{11}	0.0108	(1340)	0.75						
	0.0157								
D^{21}*(1-g(1))	0.1009	(648)	0.73						
	0.0437								
D^{31}*(1-g(2))	0.0203	(266)	0.11						
	0.0435								
D^{41}*(1-g(3))	0.0865	(124)	0.59						
	0.0538								
D^{51}*(1-g(4))	-0.0532	(55)	0.09						
	0.0941								
D^{61}*(1-g(5))	-0.0094	(34)	0.22						
	0.1058								
D^{71}*(1-g(6))	0.0995	(20)	0.65						
	0.0652								
	30,502			30,003	30,269	30,393	30,448	30,482	30,502
	28,015			28,015	28,015	28,015	28,015	28,015	28,015
	2,487			1,988	2,254	2,378	2,433	2,467	2,487

The dependent variable is the difference in log real wages (in 1992 dollars using the implicit GDP deflator). The independent variables are the difference in potential experience and potential experience squared, together with a set of dummy variables $D^{n,1}$ that equal 1 if the worker left his n-th job involuntarily during the previous year, and a dummy variable D that is the sum of all these dummy variables for all $n \geq 1$, the latter all multiplied by 1-g(n), where g(n) is the sum of the first n terms of the Harmonic series. The second column denotes the number of workers who are seen to have left their n-th job involuntarily. The third column denotes the probability that the coefficient on D^{n1} exceeds the theoretical lower bound computed in the text, which equals 0 for n = 1 and 0.0737 for $n \geq 2$. The last six columns regress the difference in log wages on the product of the combined dummy variable D and (1-g(n)) for different values of n.

Figure 1: Summary Statistics for n

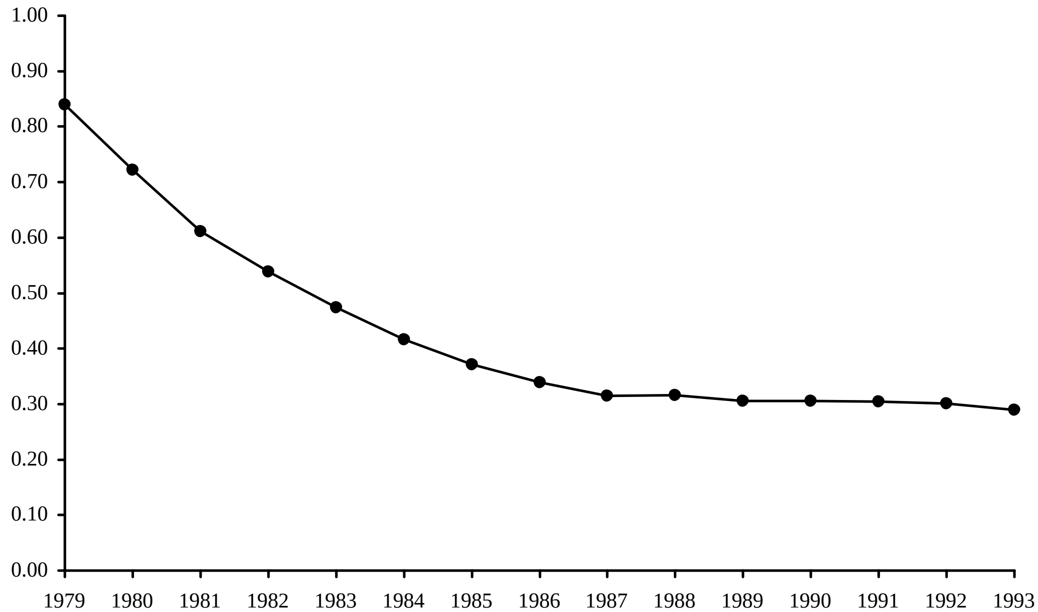


Figure 1a: Proportion of observations where no value for n was assigned

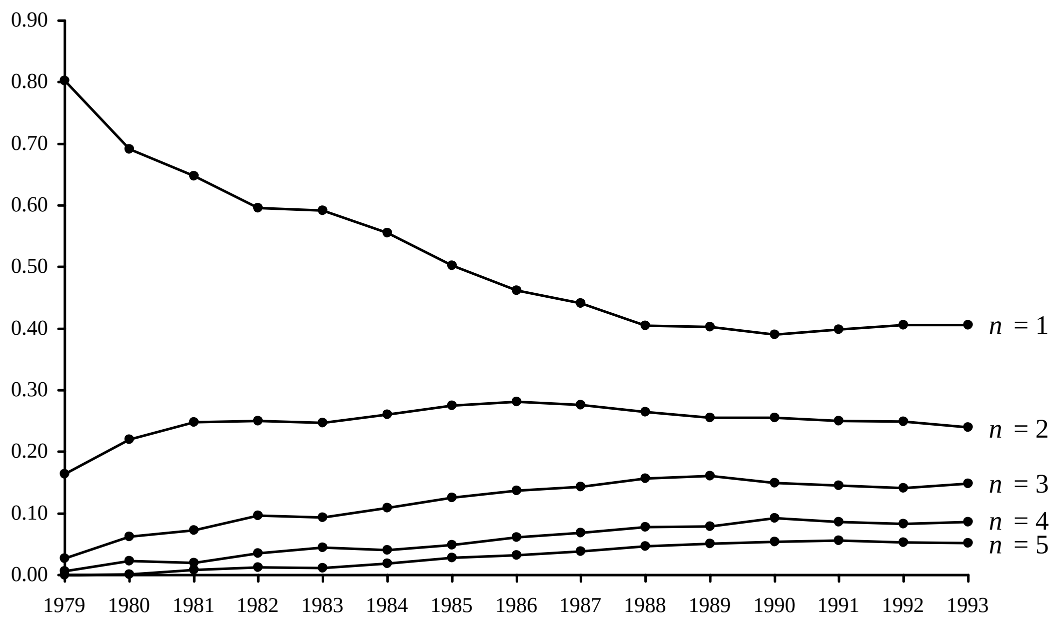


Figure 1b: Share of all observations with $n \geq 1$ for each level of n

Figure 2: Distribution of Education for
Voluntary Job Changers

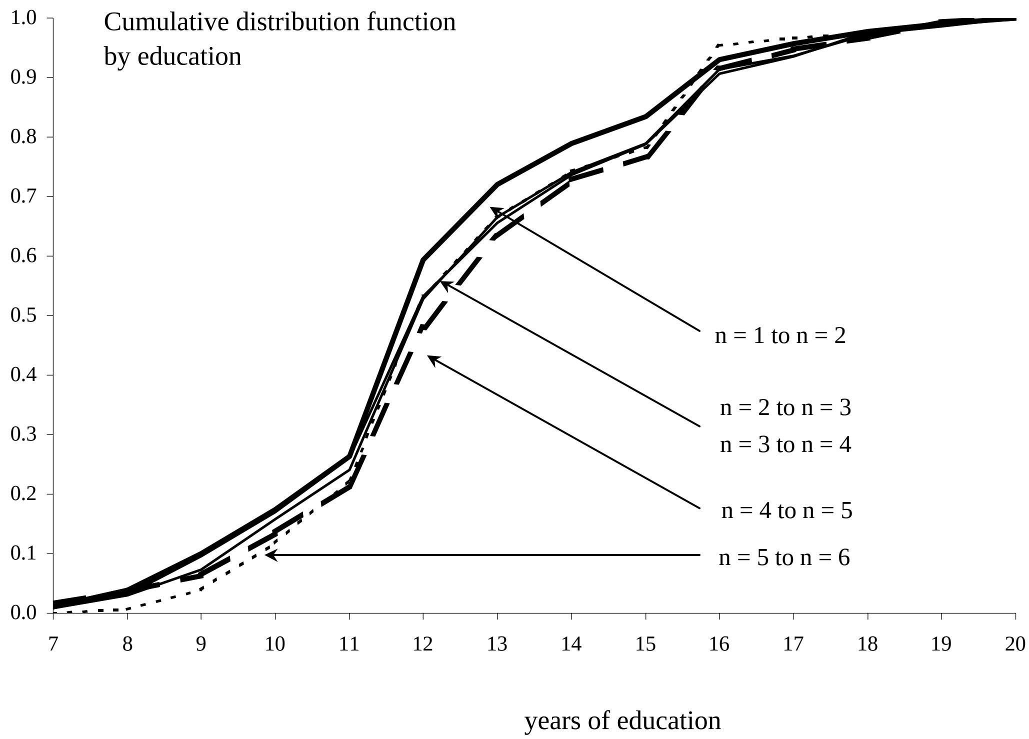
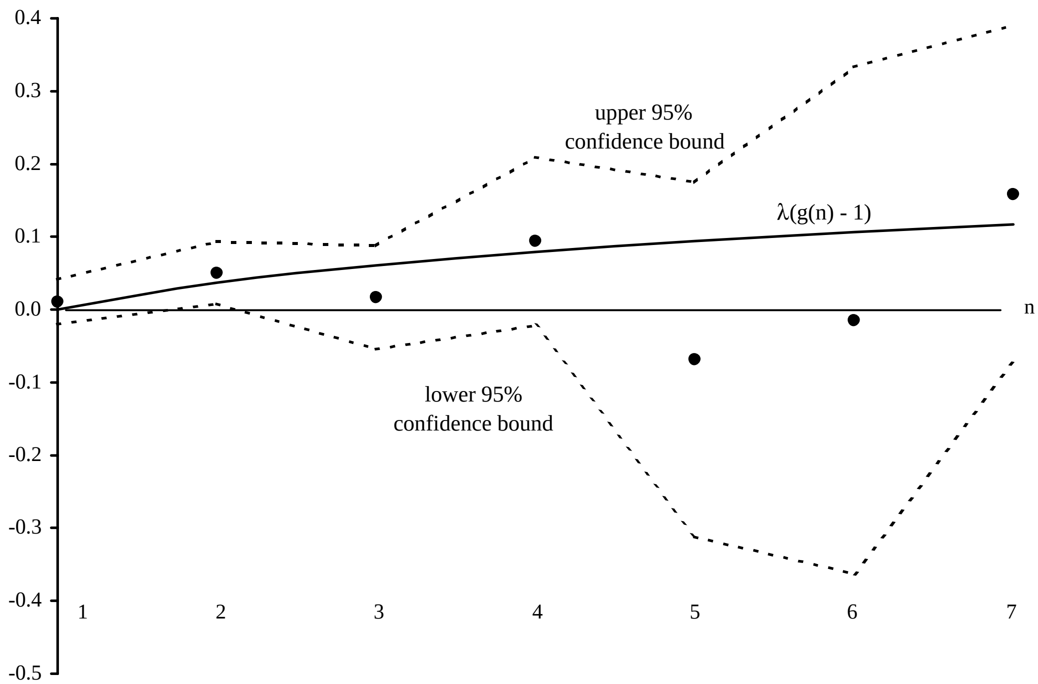


Figure 3: Change in Match Quality for Involuntary Job Changers



References

[1] Abowd, John and David Card, 1989. "On the Covariance Structure of Earnings and Hours Changes" *Econometrica*, 57(2), March, p411-45.

[2] Abraham, Katherine and Henry Farber, 1987. "Job Duration, Seniority, and Earnings" *American Economic Review*, 77(3), June, p278-97.

[3] Altonji, Joseph and Robert Shakotko, 1987. "Do Wages Rise with Job Seniority?" *Review of Economic Studies*, July, 54(3), p437-59.

[4] Altonji, Joseph and Nicolas Williams, 1997. "Do Wages Rise with Job Seniority? A Reassessment" Mimeo, Northwestern University.

[5] Altonji, Joseph and Nicolas Williams, 1998. "The Effects of Labor Market Experience, Job Seniority, and Job Mobility on Wage Growth" *Research in Labor Economics*, 17, p233-276.

[6] Arnold, Barry, N. Balakrishnan and H. Nagaraja, 1992. *A First Course in Order Statistics*. New York: John Wiley and Sons.

[7] Arnold, Barry, N. Balakrishnan and H. Nagaraja, 1998. *Records*. New York: John Wiley and Sons.

[8] Barlevy, Gadi, 2002. "The Sullying Effect of Recessions" *Review of Economic Studies*, 69 (1), January, p65-96.

[9] Beaudry, Paul and John DiNardo, 1991. "The Effect of Implicit Contracts on the Movement of Wages over the Business Cycle: Evidence from Micro Data" *Journal of Political Economy*, August, 99(4), p665-88.

[10] Bunge, John and H. Nagaraja, 1991. "The Distributions of Certain Record Statistics from a Random Number of Observations" *Stochastic Processes and Their Applications*, 38, p167-83.

[11] Burdett, Kenneth and Dale Mortensen, 1998. "Wage Differentials, Employer Size, and Unemployment" *International Economic Review*, 39, p257-273.

[12] Chandler, K. N., 1952. "The Distribution and Frequency of Record Values" *Journal of the Royal Statistical Society*, Series B, 14, p220-8.

[13] Flinn, Christopher, 1986. "Wages and Job Mobility of Young Workers" *Journal of Political Economy* 94(3, Part 2), pS88-S110.

[14] Glick, Ned, 1978. "Breaking Records and Breaking Boards" *American Mathematical Monthly*, 85(1), p2-26.

[15] Gupta, R. C., 1984. "Relationships Between Order Statistics and Record Values and Some Characterization Results" *Journal of Applied Probability*, 21, p425-30.

[16] Jacobson, Louis, Robert LaLonde, and Daniel Sullivan, 1993. "Earnings Losses of Displaced Workers" *American Economic Review*, September, 83(4), p685-709.

[17] Jovanovic, Boyan, 1979. "Job Matching and the Theory of Turnover" *Journal of Political Economy*, 87(5), Part 1, October, p972-90.

[18] Keith, Kristen and Abigail McWilliams, 1995. "The Wage Effects of Cumulative Job Mobility" *Industrial and Labor Relations Review*, 49(1), October, p121-37.

[19] Kortum, Samuel, 1997. "Research, Patenting, and Technological Change" *Econometrica*, 65(6), November, p1389-1419.

[20] Lin, G. D., 1987. "On Characterizations of Distributions via Moments on Record Values" *Probability Theory and Related Fields*, 74, p479-83.

[21] McCall, Brian, 1990. "Occupational Matching: A Test of Sorts" *Journal of Political Economy*, 98(1), February, p45-69.

[22] Mortensen, Dale, 2002. "Wage Dispersion: Why are Similar Workers Paid Differently?" Manuscript, Northwestern University.

[23] Munasinghe, Lalith, Brendan O'Flaherty, and Stephan Danninger, 2001. "Globalization and the Rate of Technological Progress: What Track and Field Records Show" *Journal of Political Economy*, 109(5), October, p1132-49.

[24] Nagaraja, H. N., 1978. "On the Expected Values of Record Values" *Australian Journal of Statistics*, 19, p70-3.

[25] Nagaraja, H. N. and Gadi Barlevy, 2002. "Characterizations Using Record Moments in a Random Record Model and Applications" Mimeo, Ohio State and Northwestern.

[26] Neal, Derek, 1999. "The Complexity of Job Mobility among Young Men" *Journal of Labor Economics*, 17(2), April, p237-61.

[27] Neal, Derek and Sherwin Rosen, 2000. "Theories of the Distribution of Earnings" *Handbook of Income Distribution*, 1, Amsterdam: Elsevier Science, North-Holland, p379-427.

[28] Nevzorov, V. B. and N. Balakrishnan, 1998. "A Record of Records" *Handbook of Statistics: Order Statistics, Theory and Methods*, v16, eds. N. Balakrishnan and C. R. Rao, Amsterdam: North-Holland, p515-70.

[29] Paxson, Christina and Nachum Sicherman, 1996. "The Dynamics of Dual-Job Holding and Job Mobility" *Journal of Labor Economics*, July, 14(3), p357-93.

[30] Pfeifer, Dietmar, 1982. "Characterizations of Exponential Distributions by Independent Non-Stationary Record Increments" *Journal of Applied Probability*, 19, p127-35.

[31] Pissarides, Christopher, 2000. *Equilibrium Unemployment Theory*, Cambridge, MA: MIT Press.

[32] Ruhm, Christopher, 1991. "Are Workers Permanently Scarred by Job Displacements?" *American Economic Review*, March, 81(1), p319-24.

[33] Topel, Robert, 1991. "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority" *Journal of Political Economy*, February, 99(1), p145-76.

[34] Topel, Robert and Michael Ward, 1992. "Job Mobility and the Careers of Young Men" *Quarterly Journal of Economics*, May, 107(2), p439-79.