Estimation of a Roy/Search/Compensating Differential Model of the Labor Market (Preliminary and Incomplete)

Christopher Taber

University of Wisconsin-Madison

Rune Vejlin

Aarhus University

February 23, 2015

1 Introduction

The four most important models of post-schooling wage determination in economics are probably human capital, the Roy model, the compensating differentials model, and the search model. All four lead to wage heterogeneity. While separating human capital accumulation from the others is quite common, we know remarkably little about the relative importance and interactions of the other three sources of inequality. The goal of this paper is to quantify the effect of each of these explanations on overall wage inequality and also to investigate how they interact.

In a human capital model, workers who have accumulated more human capital while working will earn more money than others leading to earnings inequality. The key aspect of the Roy model is comparative advantage in which some workers earn more than others as a result of different skill levels at labor market entry. Workers choose the job for which they achieve the highest level of wages. By contrast, in a compensating wage differentials model a worker is willing to be paid less in order to work on a job that they enjoy more. Thus, workers with identical skills and job opportunities can earn different salaries.Finally, workers may just have had poor luck in finding their ideal job. This type of search friction can also lead to heterogeneity in wages as some workers may work for higher wage firms. In short, one worker may earn more than another a) because he has accumulated more human capital while working (human capital), b) because he has more talent at labor market entry (Roy model), c) because he has chosen more unpleasant job (compensating differentials), or d) because he has had better luck in finding a good job (search frictions). The goal of this work is to uncover the contribution of these different components to overall wages inequality.

We develop and estimate a structural model of wage determination that contains elements of all four models. The model is estimated on Danish matched employer-employee data. We use the estimated parameters to decompose overall wage inequality into the four components in various ways. We find that while all four models are important contributors to overall earnings inequality, the Roy model inequality is the most important. The precise way in which the components matter depends on the way we measure them as there are interesting interactions between the components.

We briefly discuss the relationship between this work and the previous literature in Section 2. We then describe the model and the decomposition in Section 3. Identification is discussed in 4 and the specific econometric specification is presented in Section 5. Obtaining the right data is crucial to this exercise. Ideally one needs matched employer/employee data as well as a long panel on workers. We describe the data in section 6. Section 7 presents the auxiliary model that we use and Section 8 presents the results.

2 Relation to Other Work

Clearly there is a huge amount of work on search models, on the Roy model, on human capital acquired on the job and on compensating differentials. A full review of all of these literatures is beyond the scope of this paper. However, we discuss the relationship between our work and a few other key papers. Two important related literatures were started by Abowd, Kramarz, and Margolis (1999) and by Postel-Vinay and Robin (2002). Both of these other papers use the Declarations Annuelles des Donnees Sociales (DADS) data set which is panel data on both firms and workers from France. Abowd, Kramarz, and Margolis (1999) use a fixed effect approach to estimate firm effects and worker effects while Postel-Vinay and Robin (2002) estimate a structural equilibrium search model.

Abowd, Kramarz, and Margolis (1999) estimate a model analogous to

$$\log(W_{itj}) = X_{it}'\beta + \theta_i + \mu_j + \zeta_{ijt} \tag{1}$$

where i indexes an individual, j indexes a firm, and t indexes time. θ_i is an individual fixed effect, μ_j is a firm fixed effect, and ζ_{ijt} is independent and identically distributed. We use this as a motivation for our auxiliary model. One major way in which we will build on their work is the inclusion of compensating differentials. In estimating both the firm specific effect in wages and the firm specific effect in utility, we can simulate how much of the differences in firms wages that we observe seems to occur from market inefficiencies (search) and how much occurs as a result of workers choice (compensating differentials). Furthermore, we write down a sample selection model for our analogue of ζ_{itj} which makes interpretation of the model easier. Finally, we have all of the advantages of a structural model which further helps in interpretation of the model and allows us to use it for policy simulations.

Postel-Vinay and Robin (2002) decompose wage inequality into a search component and an ability related component. They find that for skilled workers the individual component is moderately important (close to 40%), but that for low skill workers virtually all of the inequality can be assigned to search frictions. The main components that we add is that we allow for non pecuniary benefits and comparative advantage in jobs. Postel-Vinay and Robin allow for absolute advantage only - ability is one dimension and the relative productivity of two workers does not vary across firms. By contrast, we allow for a match specific effect

(v_{ij}) meaning that some workers match better with some firms. Furthermore, our estimation and identification strategy are very different. They estimate assuming that the model is in steady state, so the information we get from looking at the rate at which people switch jobs and from the revealed preference argument is not a source of identification their model.

Bagger, Fontaine, Postel-Vinay, and Robin (2011) extends Postel-Vinay and Robin (2002) in order to incorporate general human capital accumulated while working. Their goal is to separate the concave life-cycle wage profile into search and human capital. Human capital accumulation is found to be the most important source of wage growth for workers early in their careers. However, this is soon surpassed by search-induced wage growth. This is especially true for low educated workers. In general high educated workers have a higher return to experience than low educated workers.

Another important related literature is one started by the work of Keane and Wolpin (1997) and the many papers building on their model. They estimate a model that includes compensating differentials, human capital, and Roy model inequality. They do not explicitly incorporate search frictions and do not make use of firm level data. Becker (2009) uses a framework similar to ours in that it incorporates compensating differentials into a search models. However, it does not allow for as much Roy flexibility as we do and is not estimated using firm level data and focuses more on unemployment insurance. Another nice paper that has aspects of the four models we discuss above is Sullivan (2010). This paper includes elements of all four of our models above, though this is not the main focus (and while there are search frictions it is not a standard search model). From a modeling stand point the main difference between our paper and his is the determination of wages which he specifies using a reduced form. The second main difference is the type of data we use. We show in our identification section that matched worker/firm data is essential to perform our exercise. The combination of these two components is essential for our model in distinguishing the extent to which firm/worker effects are due to Roy model inequality or compensating differentials. It is also important for distinguishing between search frictions and compensating differentials in explaining firm wage premiums.

Dey and Flinn (2005) and Dey and Flinn (2008) estimate search models with a particular type of non-wage characteristics: health insurance. Other than that, our models are quite different. Sullivan and Too (2011) is more similar in that they estimate a job search model with a general form of non-wage job characteristics. Thus, the model include search and compensating differentials. However, many differences exist between their paper and ours.

First, they specify output as being only match specific, while we allow workers and firms to have constant ability and productivity across matches as well as match specific. The reason for this choice by Sullivan and Too (2011) is likely motivated by the use of NLSY data in the estimation. The NLSY follows only workers, so it is not suited for dealing with firms. Secondly, their model is only a partial equilibrium model in the sense that workers draw a wage and a non-wage component, but there is no negotiation between firms and workers. E.g. a firm does not try to negotiate the wage down in a match where the worker have a high value of the non-wage component. Finally, their model does not include human capital.

For a good discussion of the Roy model see Roy (1951), Heckman and Sedlacek (1985), Heckman and Honoré (1990), or Heckman and Taber (2008). Rosen (1987) provides an excellent discussion of compensating differentials models and Eckstein and Van den Berg (2007) provide a nice discussion of empirical search models.

3 The Model and Decomposition

We present a continuous time model in which agents are infinitely lived. Wages are determined similarly to Cahuc, Postel-Vinay, and Robin (2006), Dey and Flinn (2005), and Bagger, Fontaine, Postel-Vinay, and Robin (2011).[1] A substantive difference between our paper and most of the search literature is that we assume that there are finite types of establishments indexed $j = 1, ..., J$ with $j = 0$ denoting nonemployment. Job offers from each type of establishment arrive at rate λ_j^n for nonemployed workers and λ_j^e for employed workers. Matches are destroyed exogenously at rate δ. After a job is destroyed we allow some individuals to immediately receive offers. Specifically, with probability P^* the worker immediately receives an offer drawn from the same distribution as for unemployed workers and can either reject or accept it.

Human capital is completely general and takes on a discrete set of values $\psi_0, ..., \psi_H$. When individuals are employed, human capital appreciates randomly to the next level (ψ_h to ψ_{h+1}) at rate λ_h and does not accumulate when people are not working. We let $\pi_{ij}\psi_h$ be the productivity of worker i at establishment j when the worker has human capital ψ_h. In the bargaining protocol presented later, the object of negotiation is the human capital rental rate. That is the employer and worker agree on a rental rate R that is fixed until

[1] Note, that we assume bargaining over wages as opposed to wage posting. Hall and Krueger (2012) show that there is mixed evidence regarding the wage determination process. In their survey around one third of all workers report having bargained over their wage. Another third reports that they had precise information about the wage before meeting the employer, which is a sign of wage posting.

the next negotiation. This means that when human capital is augmented the wage is not renegotiated but automatically rises from $R\psi_h$ to $R\psi_{h+1}$. Thus, the maximum rental rate that a firm is willing to pay is π_{ij}.[2] This means that as human capital augments, wages increase proportionally.

We do not allow for borrowin or lending. A worker i's flow utility from working at job j with wage W is $u_{ij}(W)$. The fact that this depends on j is an important part of our story that will accomadate compensating differentials-workers care about jobs above and beyond the wage that they earn.

Following Bagger, Fontaine, Postel-Vinay, and Robin (2011), a key aspect of this model is that when a worker receives an outside offer then the human capital rental rate is determined by a form of generalized Nash Bargaining between the two firms. This form of wage setting leads to efficient turnover. The rental rate is kept fixed until both parties agree to renegotiate it, or when the job spell ends.[3]

Define $V_{ijh}(R)$ to be the value function for worker i with the rental rate R working in job j and having human capital level h. Workers who are nonemployed will have flow utility U_{i0h} and value function V_{i0h}. We will let V^*_{i0h} denote the value function immediately after a match is destroyed. The difference between V^*_{i0h} and V_{i0h} is that the former incorporates the possiblity of receiving an offer immediately.

If a non-employed worker i with human capital ψ_h receives an offer from firm j, they will take the job if $V_{ijh}(\pi_{ij}) > V_{i0h}$. In general the rental rate is denoted $R_{ij\ell h}$, since it depends on worker i, the current establishment j, the best outside option ℓ, and units of human capital, h, *at the negotiation time*. For any value of R such that $V_{i0h} \leq V_{ijh}(R) \leq V_{ijh}(\pi_{ij})$ both the worker and the firm would prefer a negotiated rate of R rather than to not form the match.

An issue is that there are many such values of R. We assume that the negotiated rental rate, R_{ij0h}, for a nonemployed worker meeting firm j is defined by

$$V_{ijh}(R_{ij0h}) = \beta V_{ijh}(\pi_{ij}) + (1-\beta)V_{i0h} \tag{2}$$

where β is the worker's bargaining power.[4] Since the bargaining position for a worker who has just been laid off is nonemployment, this will also be the wage for workers who experience

[2]We are implicitly assuming the value of a vacancy is zero. One could easily relax this assumption and just redefine π_{ij} to be the maximum rental rate a firm would ever offer.

[3]In the presentation of this model we assume that workers would never want to renegotiate when their human capital augments. This does not have to be true as a worker may prefer nonemployment to the current job under some circumstances. However, this will not happen in our empirical specification so we abstract from it here.

[4]We do not derive this from a bargaining game but it as a functional form assumption: wages are indeterminate and this will give a wage that both parties will agree to. It has the nice property that if there

a match destruction but then are immediately hired to a new firm. Note that when $\beta = 1$ the worker has all of the bargaining power and extracts full rent $R_{ij0h} = \pi_{ij}$. When $\beta = 0$ the firm has all of the bargaining power and pays a value of R_{ij0h} that makes him indifferent between accepting the offer or staying non-employed.

Now suppose that worker i is working at establishment j and receives an outside offer from establishment ℓ. As in (Postel-Vinay and Robin 2002), one of three things can happen. First, the new job offer could dominate the old one, $V_{i\ell h}(\pi_{i\ell}) > V_{ijh}(\pi_{ij})$. In this case the worker will switch to the new establishment. Her new rental rate $R_{i\ell jh}$ will be determined by

$$V_{i\ell h}(R_{i\ell jh}) = \beta V_{i\ell h}(\pi_{i\ell}) + (1-\beta)V_{ijh}(\pi_{ij}) \tag{3}$$

If $V_{i\ell h}(\pi_{i\ell}) < V_{ijh}(\pi_{ij})$ then the worker has the option to renegotiate their wage. If they choose to renegotiate, their new rental rate will be determined by

$$V_{ijh}(R_{ij\ell h}) = \beta V_{ijh}(\pi_{ij}) + (1-\beta)V_{i\ell h}(\pi_{i\ell}). \tag{4}$$

Thus the worker's decision to renegotiate will just depend on whether $R_{ij\ell h}$ is higher than her current rental rate. If so, she will renegotiate the contract, if not she will ignore the offer. This condition is relatively straight forward to check. Let $R_{ij\ell h}$ be the current rental rate. If $V_{ijh}(R_{ij\ell h}) < V_{i\ell h}(\pi_{i\ell})$, then the worker will want to renegotiate. Otherwise she won't. Note that we have been a bit sloppy with notation as we use the notation $R_{ij\ell h}$ to denote the rental rate that worker i with human capital ψ_h at the time of negotiation would receive from firm j when their outside option was firm ℓ. As one can see from equations (3) and (4), it will be the same regardless of whether they started at firm ℓ and moved to j or if they started at j and then used an outside offer from firm ℓ to renegotiate their wage. Thus this is a result rather than an assumption.

To solve the model we need to be able to calculate the value functions $V_{ijt}(R)$ and V_{i0t}.

It is convenient to define

$$\Lambda_{ijh}(R) \equiv \sum_{\{\ell : V_{ijh}(R) < V_{i\ell h}(\pi_{i\ell})\}} \lambda^e_\ell$$

$$\Lambda^n_{i0t} \equiv \sum_{\{\ell : V_{i\ell h}(\pi_{i\ell}) > V_{i0h}\}} \lambda^n_\ell$$

the first of these is the sum of arrival rates that will lead to some reaction-either renogtiation or switching job. Thus for worker i with human capital h who is currently employed at firm

is surplus in the match, when $\beta = 1$ the worker get all of the surplus, when $\beta = 0$ the firm gets all of the surplus, and when $0 < \beta < 1$ the surplus between them is split.

j with rental rate R this is the arrival rate of some outside offer that will change behavior. The second item is the anaogue for nonempoyed workers-that it is the rate of arrival of an acceptable job.

We can write the value function for worker i with human capital h who is currently employed at firm j with rental rate R as

$$
\begin{aligned}
&\left(\rho+\delta+\lambda_h+\Lambda^e_{ijh}(R)\right)V_{ijh}(R)\\
&=U_{ij}(R\psi_h)+\sum_{\{\ell:V_{ijh}(R)<V_{i\ell h}(\pi_{i\ell})\leq V_{ijh}(\pi_{ij})\}}\lambda^e_\ell\left[\beta V_{ijh}(\pi_{ij})+(1-\beta)V_{i\ell h}(\pi_{i\ell})\right]\\
&+\sum_{\{\ell:V_{i\ell h}(\pi_{i\ell})<V_{ijh}(\pi_{ij}))\}}\lambda^e_\ell\left[\beta V_{i\ell h}(\pi_{i\ell})+(1-\beta)V_{ijh}(\pi_{ij})\right]+\delta V^*_{i0h}+\lambda_h V_{ijh+1}(R).
\end{aligned}
$$

Consider the different components on the right hand side of this equation. The first, $U_{ij}(R\psi_h)$, is the flow utility that the worker receives until something happens. The second component denotes outside offers that will lead the worker to renegotiate their wage but ultimately stay at the current firm. The component $V_{ijh}(R)<V_{i\ell h}(\pi_{i\ell})\leq V_{ijh}(\pi_{ij})$ defines the type of firms for which this is true. If $V_{i\ell h}(\pi_{i\ell})\leq V_{ijh}(R)$ the outside offer will not be useful for renegotiating and if $V_{i\ell h}(\pi_{i\ell})<V_{ijh}(\pi_{ij}))$ then the worker will leave to the next firm. The component in brackets represents the value function of the renegotiated wage. The next term denotes outside offers that lead the worker to leave the current firm. Again the term in brackets denotes the value function under the negotiated wage. The next two terms δV^*_{i0h} and $\lambda_h V_{ijh+1}(R)$ respectively represent the events in which the worker is laid off and in which human capital augments.

When $h=H$ we get an expression that is identical except that it no longer contains the possibility of human capital augmenting:

$$
\begin{aligned}
&\left(\rho+\delta+\Lambda^e_{ijh}(R)\right)V_{ijH}(R)\\
&=U_{ij}(R\psi_H)+\sum_{\{\ell:V_{ijH}(R)<V_{i\ell H}(\pi_{i\ell})\leq V_{ijH}(\pi_{ij})\}}\lambda^e_\ell\left[\beta V_{ijH}(\pi_{ij})+(1-\beta)V_{i\ell H}(\pi_{i\ell})\right]\\
&+\sum_{\{\ell:V_{i\ell H}(\pi_{i\ell})<V_{ijH}(\pi_{ij}))\}}\lambda^e_\ell\left[\beta V_{i\ell H}(\pi_{i\ell})+(1-\beta)V_{ijH}(\pi_{ij})\right]+\delta V^*_{i0H}.
\end{aligned}
$$

There is only one way for the status to change following a nonemployment spell-the worker can take a job. This leads to the simpler formulation

$$
\left(\rho+\Lambda^n_{i0h}\right)V_{i0h}=U_{i0h}+\sum_{\{\ell:V_{i\ell h}(\pi_{i\ell})>V_{i0h}\}}\lambda^n_j\left[\beta V_{i\ell h}(\pi_{i\ell})+(1-\beta)V_{i0h}\right].
$$

The first term is the flow utility and the second denotes the outcome in which an offer is received that dominates nonemployment. The term in brackets represents the value function under the renegotiated rate.

Finally the value function for workers immediately after their match is destroyed is

$$V_{i0}^{*}(h_t) = P^{*}\frac{\sum_{\{\ell : V_{i\ell}(\pi_{i\ell},h_t) > V_{i0}(h_t)\}} \lambda_\ell^n V_{i\ell}(W_{i\ell 0}, h_t)}{\sum_\ell \lambda_\ell^n} + \left(1 - P^{*}\frac{\Lambda_{i0t}^n}{\sum_\ell \lambda_\ell^n}\right) V_{i0}\left(h_t\right)$$

The first term is the probability of an acceptable job and the second is the result from either no offer or an unacceptable offer.

This is the full model. There are many other features in the labor market that we have abstracted from. This is intentional. Our goal here is not to write down the most complicated model that is computationally feasible, but rather to write down the simplest model that captures the essence of our four models and allows us to distinguish between them. To see this, next consider a decomposition that allows us to understand the various components. We can choose any measure of wage inequality that we want (for example the variance of log wages). In the context of the model we have written down, one can see the different sources of wage inequality:

- Worker variation in potential rental rates $\pi_{i1}, ..., \pi_{iJ}$ leads to "Roy model" inequality
- Variation in the function $U_{ij}(\cdot)$ across workers accommodate "compensating differentials" inequality (and the mean of $U_{ij}(\cdot)$ will vary across jobs)
- λ_j^e and λ_j^n incorporate search frictions - note that these effect wage inequality in two different ways: directly through the job at which one takes, and indirectly through the negotiation process
- Variation in ψ_h incorporates human capital

After estimating the parameters of the model, we use it to decompose overall post-schooling wage inequality into the different components. An orthogonal decomposition does not exist, so one can perform this decomposition in a number of different way. More generally the different sources interact. Perhaps most importantly, the fact that workers have comparative advantage in some jobs rather than others interacts with search frictions since search frictions restrict not only access to firms with generous wage policies but also to firms that are good matches.

Thus, many possible ways to decompose wages exist and we use different ones to highlight different features of the model. The following simulations represent one example of a decomposition we take. We sequentially take the following steps.

- a) First simulate the cross section variance of wages using all parameters (which should be approximately the same as overall wage inequality in the data).
- b) eliminate variation from human capital. We can do this by setting λ_h arbitrarily high so that workers obtain their maximum human capital immediately.
- c) Eliminate variation coming from heterogeneity in firm monoposony power through renegotiated wages. We can do this by setting $\beta = 1$ so that workers collect their productivity immetiate.y.
- d) Eliminate heterogeneity coming from worker luck in the job offers they have received. We do this by setting λ_j^e arbitrarily large so that employed workers move to the preferred job immediately.
- e) Eliminate Roy model inequality. To do this we set rental rates to be $R_{ij} = E(R_{ij} \mid j)$ for all j eliminating variation due to skill differences so that firms pay constant wages (but holding preferences across jobs constant).
- f) Eliminate inequality coming out of choice of jobs. We can do this by assuming individuals choose jobs to maximize wages only.

The difference between a) and b) is due to human capital , the difference between b) and d) is due search frictions, the difference between e) and d) is due to Roy model inequality, and the remaining fraction in e) is due to compensating differentials. After implementing f) there is nothing left.

4 Identification

In this section we discuss non-parametric identification of our model. The details are given in Appendix A. We show which aspects of the model can and cannot be identified. We view both parts as important. This is empirically relevant in that we can not credibly simulate counterfactuals that are not identified from the data. We will respect this in our counterfactual exercises below by only simulating counterfactuals that can be identified.

Specifically, it will turn out that two aspects of our model are not identified. First, we show that one can not hope to identify R_{ij} (the rental rate that worker i would get at job j) for a worker i on a job j that they would never take (i.e. $V_{ij} < V_{i0}$). While obvious at some level, it is important to keep in mind that this limits the type of counterfactuals which can be simulated. We view this not as a limit of our model or particular data, but rather as a fundamental identification problem that will be an issue for non-parametric identification with any model and data set. It is essentially the identification at infinite problem discussed in (Heckman 1990) and (Heckman and Honoré 1990). Without an exclusion restriction moving the conditional probabiliy of doing each job to one, we can not identify the full unconditional distribution of wages in that job.

Second, we can not nonparametrically identify the bargaining parameter β. The reason is that we can use a revealed preference argument to identify the preference workers have across jobs, but we can not identify the levels of utility. We can, however, test whether $\beta = 1$ and simulate a model without bargaining in which the worker receives a competitive wage offer ($\beta = 1$).

Identification crucially depends on two types of arguments that are somewhat nonstandard and also depends on some special aspects of our data. The first is that we follow Villanueva (2007) (and others) by using a revealed preference argument that a worker has shown a preference for one job over another if they directly leave the first job to start the second. As a result it will be very important for us to distinguish job-to-job transitions (in which we will use our revealed preference argument) from job-to nonemployment-to job transitions (where we are not willing to use this argument). Intuitively, if we consistently observed that workers were willing to take wage cuts to go to a certain firm, this would indicate that this firm had high non-pecuniary benefits. A limitation of this approach is that in practice not every job-to-job transition is voluntary. In the model and empirical work we allow for some job to job transitions to be involunary through P^* but we abstract from this possibility here and assume $P^* = 0$.

A second key aspect of the model is that we show that the arrival rates (λ_j^e) can be identified by the rate at which workers switch jobs. If the reason that all workers do not work for the highest paying firm is because of search frictions, then eventually they should match with the highest paying firm. Thus, if search frictions are a very important component of inequality, the rate of job switching should be fairly slow. However, if search frictions are relatively unimportant (arrival rates are high) workers will quickly receive an offer from their preferred firm.[5] Thus it is important to have high quality panel data on job switching and also matched employer-employee data.

Proving general nonparametric identification of the model when J is very large seems overly tedious, so instead we focus on a simpler case to illustrate how our model can be identified. There are two different types of jobs workers can get A and B and human capital only takes on two values, $h = \{0, 1\}$. We fully expect our result to generalize to larger (but finite) J.

In the appendix we first consider what can be identified without data on wages. We assume that we observe workers from time 0 to T and that all workers begin their working life nonemployed.[6]

In the model we can partition people into types depending on their preferences for jobs and nonemployment conditional on education. We show that from the observed job choices and the timing of movements we can identify $\delta, \lambda_A^n, \lambda_B^n, \lambda_A^e, \lambda_B^e$ and the proportions of each different types.

We next incorporate information from wages. Let w_{it}^m denote the log of the wage measured at time t if the worker is working at time t. To incorporate the main features of the Danish data, we assume that we only observe wages at a finite number of times. For simplicity assume it is at the integers ($t = 1.0, 2.0, ...$). Without loss of generality we can normalize the initial level of human capital $\psi_0 = 1$. In the appendix we show that we can identify the wage distribution for each type in each state of the world that we might observe that type.

For example if the type is someone who prefers job B to job A and job A to nonemployment, for each level of human capital there are six possible labor market statuses:

1. they were hired into and A firm and have yet to receive an outside offers
2. they were hired into an A firm and received an outside offer from another A firm
3. they were originally hired into an A firm and then were poached by a B firm
4. they were hired into a B firm with no outside offer
5. they were hired into a B firm with an outside offer from an A firm

[5] It is important to point out that this does not necessarily incorporate all forms of search frictions. If the worker's first job restricts all jobs they can subsequently obtain, this will look identical to what we call Roy model heterogeneity. However, the type of search friction we have identified in the text of this paper is the most common type in equilibrium search models such as Burdett and Mortensen (1998) or Postel-Vinay and Robin (2002).

[6] In the data this is not true for all workers. However in terms of identification we can focus only on those workers who we observe at the beginning of their careers.

6. they were hired into a B firm with an outside offer from another B firm

We show that we can identify the joint distribution of the wages in each of these six situations conditional individuals of this type.

Other types are simpler. For example consider the types of individuals that prefer job B to job A but prefer nonemployment to job A. They have only two labor market statuses:

1. they were hired into a B firm with no outside offer (or equivalently an irrelevant offer from a type A firm)

2. they were hired into a B firm with an outside offer from another B firm

We can identify the joint distrubution of wages for these types. However, we can say nothing about the wage they would receive from an A firm because the data is completely uninformative about this possibility.

Next consider identification of the bargaining parameter β. In the appendix, we explicitly show that it is not identified. The issue can be seen by considering the case above with workers that would only take job B. There are only two relevant states of the world for them. In the second state of the world in which two B firms have bid for them, they will receive a wage equal to their productivity at firm B which is π_{iB}. In the other state of the world they bargain over their wage so β is important in determining their wage in this state. However we can not identify it without knowing their intensity of preferences which is fundamentally unidentified from revealed preference. For example, consider a worker who receives a wage that is very close to π_{iB}. That might be because β is very close to 1 (i.e. the worker has a lot of bargaining power) or it could be because the workers is almost indiffernt between nonemployment and working so that their reservation wage is very close to π_{iB}. As long as $\beta < 1$ for any β we can find a reservation utility to reconcile the observed wage. In the empirical model we fix this issue not by setting β ex-ante but in normalizing the intensity of preferences.

It is important to point out that even though we can not identify β nonparametrically, we can perform two interesting counterfactuals. Completely eliminating search friction will lead people to be paid their productivity at their preferred job-this is identified. We will also simulate a counterfactual where we get rid of inequality arising from the bargaining process (the difference between the wage in status1 and status 2). We can do that by setting $\beta = 1$ and this counterfactual is identified because the worker will be paid π_{ij} in that case.

5 Econometric Specification/Parameterization

We assume that there are a large number of people in the economy but a finite number of J employer types. Our econometric specification differs from the identification section above in that we do not observe establishment types. We will assume that multiple establishments in the data will be of the same type but we will not know exactly which those are. The key aspect from the data is that workers who work for the same establishment also work for the same establishment type. We will let $j_i(t)$ denote the job (establishment) held by worker i at time t. We assume that for each worker we get to observe them from when they enter the labor market until time period T_i. We also observe M_i different wages for workers at times $t = t_1, ..., t_{M_i}$.

To keep the idea of the data simple assume that every moment $t \in [0, T_i]$ we observe the firm for which the worker worked (call this $j_i(t)$) and without loss of generality let $j_i(t) = 0$ denote non-employment.

The transition parameters δ, λ^e, and λ^n take the same form as in the earlier sections though they do not vary across jobs although we allow δ to vary across individuals

$$\log(\delta_i) = d_0 + \zeta_i$$

with $\zeta \sim N\left(0, \sigma_\zeta^2\right)$. We let P^* be the probability that the individual receives an offer immediately after being laid off.

The maximum offered rental rate is specified as

$$\log(\pi_{ij}) \quad = \quad \theta_i + \mu_j^w + v_{ij}^w.$$

Again the easiest way to think about this is as the marginal productivity of a worker and that firms have constant returns to scale and no capacity constraints. This implies that hiring this worker does not restrict hiring of any other workers, nor does it affect their marginal productivities. We only get to observe wages with normal measurement error ξ_{it}, with mean zero and variance σ_ξ^2.

The tastes for jobs are

$$u_{ij}(W) = \alpha \log(W) + \mu_j^u + v_{ij}^u.$$

Note that all that matters for job to job turnover is

$$\begin{aligned} u_{ij}(\pi_{ij}\psi_t) =& \alpha\left(\theta_i + \mu_j^w + v_{ij}^w\right) + \alpha \log(h_t) + \mu_j^u + v_{ij}^u \\ =& \alpha\theta_i + \alpha \log(h_t) + \left(\alpha\mu_j^w + \mu_j^u\right) + \left(\alpha v_{ij}^w + v_{ij}^u\right). \end{aligned}$$

We assume that the joint distribution of (u_j^w, u_j^u), the joint distribution of (v_{ij}^w, v_{ij}^u), the distribution of θ_i, and the distribution of ξ_{it} are all independent of each other. Thus, (v_{ij}^w and v_{ij}^u) can be interpreted as match-specific variation around the mean. Furthermore we assume that ξ_{it} is i.i.d. across time. We normalize $var(v_{ij}^u) = 1$. Moreover, we assume that $\left(v_{ij}^u, v_{ij}^w\right)$ are jointly normal which gives us two parameters, $cov\left(v_{ij}^u, v_{ij}^w\right)$ and $var\left(v_{ij}^w\right)$. However, as should be clear from the equation immediately above we can not separately identify $cov\left(v_{ij}^u, v_{ij}^w\right)$ from α, so we normalize $cov\left(v_{ij}^u, v_{ij}^w\right) = 0$ and estimate α along with σ_{v^w}, the standard deviation of v_{ij}^w.[7] Notice, that the covariance of (u_j^w, u_j^u) is left unrestricted, so the common part of productivity and non-pecuniary returns are allowed to be correlated.

We also assume that θ_i is normal with mean E_θ and variance σ_θ^e.

Human capital evolves as

$$\log(\psi_h) = b_1 h + b_2 h^2 + b_3 h^3$$

We use a cubic spline so there are two free parameters and we choose the third to impose that

$$\frac{\partial \log(\psi_H)}{\partial h} = 0$$

We fix λ_h=1.

We take a very simple specification for the value of nonemployment by assuming

$$u_{i0t} = \alpha E(\theta_i) + \gamma_\theta \left(\theta_i - E(\theta_i)\right) + \nu_{io}^n$$

with $\nu_{i0}^n \sim N(0, \sigma_\nu^2)$.

We tried to choose a relatively parsimonious way to approximate the distribution of $\left(\mu_j^u, \mu_j^w\right)$ which is a discrete distribution. With no obvious parametric alternative we use the following one:

$$\begin{aligned}
\mu_j^u &= f_1 \left[U_1(j) + f_3 U_2(j)\right] \\
\mu_j^w &= f_2 \left[f_3 U_1(j) + U_2(j)\right]
\end{aligned}$$

where $U_1(j)$ and $U_2(j)$ are distributed as discrete uniform across [-1,1]. In our specification we allow each of U_1 and U_2 to take ten different values and assume these are unrelated to

[7]This is a normalization to the extent that it is innocuous in terms of the model we estimated since choices across jobs depends only upon the overall $u_{ij}(\pi_{ij} h_t)$ not on its components. However, it is not a normalization in that in counterfactuals in which π_{ij} changes, the preference across jobs depends on this assumption. When we use it in this way it assumes that the reason why workers tend to prefer to work at a high productive matches is because they receive high wages at those jobs whether than because they tend to like the non-wage characteristics of jobs on which they are particularly productive. This seems like a very reasonable assumption to us.

each-other giving us one hundred different firm types. Essentially f_1 governs the variance of μ_j^u, f_2 governs the covariance of μ_j^2, and f_3 governs their correlation.

We fix $\rho = 0.05$. This leaves a total of 18 parameters:

$$\left[d_0, \lambda^n, \lambda^e, E_\theta, \sigma_\theta, \sigma_\xi, \sigma_{v^w}, \alpha, f_1, f_2, f_3, b_1, b_2, \beta, P^*, \sigma_\zeta^2, \sigma_\nu^2, \gamma_\theta\right].$$

Indentification and Functional Forms

Before we proceed we provide some discussion and why choose the normalizations that we do. First, as is well known from (Flinn and Heckman 1982) it is generally not possible to separate the reservation wage from the arrival rate of jobs. In general one could allow for full heterogeneity both in arrival rates λ^n and λ^e as well as reservation utility u_{i0t}. Of course the (Flinn and Heckman 1982) result applies to heterogeneity as well-we can distinguish heterogeneity in reservation utility from heterogeneit in arrival rates. While our model might be formally identified by functional form restricitions, we do not try to separately idenify these effects. We deal with this problem by allowing heterogeneity in reservation utility, u_{i0t}, only (through ν_{io}^n) so treating λ^n and λ^e as the same for all individuals. That still leaves the problem that we can not separate the level of λ^n from the level of reservation utility. We fix this problem by fixing the level of reservation utility so that a person with average ability ($\theta_i = E(\theta_i)$) and average tastes for leisure ($\nu_{i0}^n = 0$) will have a reservation utility of $\alpha E(\theta_i)$ which means that when they have no human capital their acceptance rate of jobs from nonemployment would be roughly 50%. It is also important to recognize that once we have imposed these restrictions, what is relevant for decisions is the combination of λ^n, λ^e, and u_{i0t} making interpretation of these parameters separately from eachother difficult.

The second issue is the scale of preferences. As is well known from discrete choice models the scale of the utility is not identified. We fix this problem by normalizing $var(v_{ij}^u) = 1$. Recall from the discussion in the identification section that β is not identified. The issue there is that it was impossible to distinguish bargaining power from the intensity of preferences because all we could identify ordinal utility not cardinal utility. Once we have normalized $var(v_{ij}^u) = 1$, we no longer have this problem and β is identified essentially from the importance of the bargaining process. The larger is β the smaller will be the importance of this process. Note as well that since β is identified from this normalization it can not be interpreted literally and can not be compared across different samples.

6 Data

We use Danish register data of two types. The first is weekly spell data that covers all individuals aged 15-70 in Denmark from 1985 to 2003. The data is constructed using various sources of register data, see Bunzel (2010) and Bobbio (2010) for a longer description. The data generated from these sources consists of a worker identifier, a firm and establishment identifier, start and end date of the spell, and a state variable. The states are employed, unemployed, self-employed, retirement, and non-participation. Non-participation is a residual state in the sense that it means that we do not observe the worker in any of the available registers. The second type of data is annual cross-section data from the Danish register-based matched employer-employee data set IDA.[8] IDA contains socioeconomic information on workers and background information on employers, and covers the entire Danish population age 15 to 70. IDA contains the annual average hourly wage for the job occupied in the last week of November.

We choose to identify employers at the establishment level. Thus, the unit of observation is a worker, year, state, establishment. We differ on this point from most of the empirical search literature which uses firms as the employer unit. However, using establishments have three advantages in the current setting. The employer identifier is not well defined over time since firms might change the legal unit (and hence firm identifier) without changing anything else. The establishment identifier is consistent over time.[9] Secondly, when thinking about compensating differentials the most appropriate unit seems to us to be the establishment and not the legal firm. The third advantage is that we are able to break up larger firms into separate establishments. Treating all workers within the same large firm the same seems to be inappropriate. This is especially important in the government sector, where firms tend to be large and potentially cover many different types of establishments. While this is still not a completely satisfactory way of dealing with the government sector, we think it is much better than treating the government sector as a single firm or dropping it from the data.

6.1 Sample Selection Criteria

We use the following sample selection criteria[10]

[8] Integreret Database for Arbejdsmarkedsforskning (Integrated Database for Labor Market Research) is constructed and maintained by Statistics Denmark.

[9] The establishment is constructed by Statistics Denmark and is the same across years if one of three criteria is met: same owner and industry; same owner and workforce; same workforce and either same address or same industry.

[10] See Appendix for a more thorough description.

- We censor workers after age 55 and disregard spells before labor market entry (or age 19) defined as the time of highest completed education (and not observed in education later).
- We disregard workers with errors in educational information.
- Temporary non-employment spells shorter than 13 weeks are deleted and non-employment shorter than 3 weeks are allocated to the last of the two employment spells.
- We censor workers when they enter self-employment or retirement.
- We delete workers that have gaps in their spell histories. This could arise if the worker for some reason have missing IDA data in a given year.

In Table 1 the effects of the data steps taken above are described for each step.

Since the model is cast in steady state and cross-sectional wages therefore have no trend we detrended wages in logs by gender-educational groups conditional on experience. We do this since the composition of workers changes over the sample period. Therefore it is important not to impose e.g. average wages to be constant over all years.

Since job-to-job transitions play a vital role for the identification of our model we will ignore transitions from two types of establishments. The first are transitions for workers to or from establishments with missing ID (0.5 percent, cf. Table 2). We will also ignore job-to-job transitions from closing establishments or establishments with mass layoffs.[11]

6.2 Descriptive Statistics

In this section we present different descriptive statistics for the sample used in this study. The number of years and the number of establishments are important for the identification of the model. Table 2 show statistics for these measures using repeated cross-sections.

The worker is on average 11 years in the sample and are employed in 2.7 different establishments. There are almost as many women as men in the sample. This is because we are not censoring or deleting public employees of which many are women. The workers have on average twelve years of education. However, this changes some over the sample period, since entering workers are better educated that those leaving the sample. The average cross-section age is 38 with the earliest labor market entry at age 19 and the highest age in the sample

[11]A mass layoff is defined as the establishment having more than 15 workers and the next year only has 30 percent or less left.

being age 55. A total of 84 percent are employed in general, while 32 percent are employed in the public sector. Finally, in a given cross-section we miss the establishment identifier for 0.5 percent of all employment observations. The average labor market experience is 13 years.

Figure 1 displays the estimate from a Kaplan Meier estimator of the survival probability for employed, unemployed, and non-employed. The unemployment spells have a much shorter duration than both employment and non-employment in general. Notice, that the survival rate for non-employment does not seem to approach zero. This could be due to the fact that the sample also include workers that are actually not in the labor force, i.e. not actively looking for a job. However, we do not view this as a problem, since in the model we allow for workers who simply choose not to take a job. Turning to the employment spells we can see that there is a slower decline in the survival probability. The probability of working in the same establishment after the initial two years is 60 percent. The dips in both the survival rates for employment and non-employment comes from a "New Year" effect. In the reported data there is an over representation of state changes at January 1st each year. We suspect that some the state changes are coming from transitions the past year that have not been reported correctly.

7 Auxiliary Model

We estimate our model using Indirect Inference (Gourieroux, Monfort, and Renault 1993). Our approach is to use the argument in the identification section as a guide to which aspects of the data we should be using to identify the different parameters. To keep the relationship between the parameters and the data as transparent as possible we focus on the exactly identified case. In particular, for each parameter we choose one auxiliary parameter that we think is useful for identifying it. While we use this language, it is not precisely how the estimation works. In practice, all the auxiliary parameters are useful for identifying all of the structural parameters. However, we find that this approach to be highly beneficial for us in understanding the mapping between the parameters and the data.

7.1 Notation

Let

- $i = 1, ..., N$ index individuals

- $\ell = 1, ..., L_i$ index employment spells-a spell of consistent employment with no non-employment in between
- $j = 1, ..., J_{i\ell}$ index a job spell that occurs within employment spell ℓ for individual i
- $t = 1, ..., T_{i\ell j}$ index the set of wage observations on job spell $i\ell j$.
- $f_{i\ell j}$ the firm associated with this job spell
- $1, ..., Q$ be the number of establishments
- $D_{i\ell j}$ the duration of time that the worker worked on job spell $i\ell j$
- $w^m_{i\ell jt}$ the tth wage observation at job $i\ell j$
- $E^m_{i\ell jt}$ the tth experience observation at job $i\ell j$
- $TE^m_{i\ell jt}$ the tth tenure observation at job $i\ell j$. It is set to 0 at the first November cross-section in job $i\ell j$ and from that it increases with $E^m_{i\ell jt}$.
- $\ell^m = 1, ..., L^m_i$ index employment spells that contains a November cross-section
- $w^m_{i\ell 11}$ be the first wage observation at employment spell ℓ^m for individual i
- $E^m_{i\ell 11}$ is the corresponding first experience observation at employment spell ℓ^m for individual i
- $k = 1, ...K_i$ the number of nonemployment spells for individual i
- D^n_{ik} the duration of nonemployment spell ik
- $k^n = 1, ...K^u_i$ the number of nonemployment spells for individual i where experience at the first November cross-section following the end of nonemployment spell is available.
- E^n_{ik} is experience at the first November cross-section following the end of nonemployment spell ik
- $C_{i\ell j}$ is an indicator taking the value one if job spell $i\ell j$ is left-censored. A spell is left censored if it is the first spell that we observe in the data.

7.2 Definition of Transition Variables

I propose we use the following for moments now. Using the same term as last time, but defined in a very different way

$$S_{i\ell j} \equiv \begin{cases} 1 & \text{if spell } i\ell j \text{starts with a JJ transition and do not end with a JJ transition} \\ -1 & \text{if spell } i\ell j \text{ends with a JJ transition and did not start with a JJ transition} \\ 0 & \text{otherwise} \end{cases}.$$

where JJ transition means a Job-to-Job transition. Notice, if the spell is left censored then we assume that it starts from non-employment. Likewise, if the job is right censored we assume that it ends in a firing. I.e. the sum of $S_{i\ell j}$ across jobs for an individual can therefore be both -1 and 1. We thus define

Define

$$\overline{S_i} = \frac{1}{\sum_{\ell=1}^{L_i} J_{i\ell}} \sum_{\ell=1}^{L_i} \sum_{j=1}^{J_{i\ell}} S_{i\ell j}$$

So $\overline{S_i}$ is the average number of job-to-job transitions per job spell for individual i.

$$\widetilde{S}_{i\ell j} \equiv \left(S_{i\ell j} - \overline{S}_i\right)$$

This way $\widetilde{S}_{i\ell j}$ should sum to zero for each individual when summing over jobs, but not when summing over firms.

For consistency reasons we will sometimes need to take the individual out of the calculation, so define:

$$\widetilde{S}_{-i\ell j} = \frac{\sum_{i^*=1}^{N} \sum_{\ell^*=1}^{L_{\iota^*}} \sum_{j^*=1}^{J_{i^*\ell^*j^*}} \widetilde{S}_{i^*\ell^*j^*} 1\left[i^* \neq i, f_{i^*\ell^*j^*} = f_{i\ell j}\right]}{\sum_{i^*=1}^{N} \sum_{\ell^*=1}^{L_{\iota^*}} \sum_{j^*=1}^{J_{i^*\ell^*j^*}} 1\left[i^* \neq i, f_{i^*\ell^*j^*} = f_{i\ell j}\right]}$$

where $1[\cdot]$ is the indicator function. In other words $\widetilde{S}_{-i\ell j}$ is just the average value of $S_{i^*\ell^*j^*}$ for people who work at firm $f_{i\ell j}$ excluding individual i.

And define the number of JJ separations, s^q_{-i}, and JJ hires, h^q_{-i}, for each firm where individual i do not contribute. Again we take transitions from closing establishments out.

$$s^q_{-i} = \sum_{i^*=1}^{N} \sum_{\ell=1}^{L} \sum_{j=1}^{J_{i\ell j}-1} 1[q = f_{i^*\ell j}, i \neq i^*]$$

$$h^q_{-i} = \sum_{i^*=1}^{N} \sum_{\ell=1}^{L} \sum_{j=2}^{J_{i\ell j}} 1[q = f_{i^*\ell j}, i \neq i^*]$$

Also define

$$h_{-i\ell j} \equiv \frac{h_{-i}^{f_{i\ell j}}}{h_{-i}^{f_{i\ell j}} + s_{-i}^{f_{i\ell j}}}$$

$$\widetilde{h}_{-i\ell j} \equiv \frac{h_{-i}^{f_{i\ell j}}}{h_{-i}^{f_{i\ell j}} + s_{-i}^{f_{i\ell j}}} - \frac{1}{\sum_{\ell^*=1}^{L_i} J_{i\ell^*}} \sum_{\ell^*=1}^{L_i} \sum_{k=1}^{J_{i\ell}} \frac{h_{-i}^{f_{i\ell^* k}}}{h_{-i}^{f_{i\ell^* k}} + s_{-i}^{f_{i\ell^* k}}}.$$

7.3 Definition of Wage Variables

The worker effect is just mean worker wage over his working life

$$\overline{w_i} = \frac{\sum_{\ell=1}^{L_i} \sum_{j=1}^{J_{i\ell}} \sum_{t=1}^{T_{i\ell j}} w_{i\ell jt}^m}{\sum_{\ell=1}^{L_i} \sum_{j=1}^{J_{i\ell}} T_{i\ell j}}$$

When $T_{i\ell j} > 0$, define

$$\overline{w_{i\ell j}} = \frac{1}{T_{i\ell j}} \sum_{t=1}^{T_{i\ell j}} w_{i\ell jt}^m,$$

and define it to be zero otherwise. The distinction between $\overline{w_i}$ and $\overline{w_{i\ell j}}$ is intentional-we will use them at different points. Then when $T_{i\ell j} > 0$we define

$$\widetilde{w}_{i\ell j} \equiv \overline{w_{i\ell j}} - \frac{\sum_{\ell^*=1}^{L_i} \sum_{j^*=1}^{J_{i\ell^*}} \overline{w_{i^*\ell^* j^*}}}{\sum_{\ell^*=1}^{L_i} \sum_{j^*=1}^{J_{i\ell^*}} 1\left[T_{i\ell^* j^*} > 0\right]}$$

This now has the nice feature that we think of as standard-it will sum to zero across jobs for each individual.

Analogous to $\widetilde{S}_{-i\ell j}$define:

$$\widetilde{w}_{-i\ell j} \equiv \frac{\sum_{i^*=1}^{N} \sum_{\ell^*=1}^{L_{\iota^*}} \sum_{j^*=1}^{J_{i^*\ell^* j^*}} \widetilde{w}_{i^*\ell^* j^*} 1\left[i^* \neq i, f_{i^*\ell^* j^*} = f_{i\ell j}\right]}{\sum_{i^*=1}^{N} \sum_{\ell^*=1}^{L_{\iota^*}} \sum_{j^*=1}^{J_{i^*\ell^* j^*}} 1\left[i^* \neq i, f_{i^*\ell^* j^*} = f_{i\ell j}\right]}$$

7.4 Definition of Experience Variables

When $T_{i\ell j} > 0$, define

$$\overline{E_{i\ell j}} = \frac{1}{T_{i\ell j}} \sum_{t=1}^{T_{i\ell j}} E_{i\ell jt}^m,$$

7.5 Moments

In general notice that a lot of the moments are calculated over different samples, since not all variables are defined for each job spell.

There are 6 parameters that are important for turnover: $\delta, \lambda^e, \lambda^n, f_1, \alpha_u^\theta, \alpha_u^h$

Here are a set of moments we could use to identify them:

- δ : Average length of employment spell:

$$\overline{L} = \frac{\sum_{i=1}^{N} \frac{\sum_{\ell=1}^{L_i}\left(\sum_{j=1}^{J_{i\ell}} D_{i\ell j}\right)}{L_i}}{\sum_{i=1}^{N} 1[L_i > 0]}$$

- λ^n-nonemployment spell. Similarly we use

$$\overline{K} = \frac{\sum_{i=1}^{N} \frac{\sum_{k=1}^{K_i} D_{ik}^n}{K_i}}{\sum_{i=1}^{N} 1[K_i > 0]}$$

- α_u^θ : Similar to above, but a covariance between duration and $\overline{w}_i$ and this only uses observations for which $\overline{w_i}$ can be calculated

$$\frac{\sum_{i=1}^{N}\sum_{k=1}^{K_i} D_{ik}^n \overline{w_i}}{\sum_{i=1}^{N} K_i} - \left(\frac{\sum_{i=1}^{N}\sum_{k=1}^{K_i} D_{ik}^n}{\sum_{i=1}^{N} K_i}\right)\left(\frac{\sum_{i=1}^{N}\sum_{k=1}^{K_i} \overline{w_i}}{\sum_{i=1}^{N} K_i}\right)$$

- α_u^h : Similar to a

$$\frac{\sum_{i=1}^{N}\sum_{k^n=1}^{K_i^n} D_{ik}^n E_{ik}^n}{\sum_{i=1}^{N} K_i^n} - \left(\frac{\sum_{i=1}^{N}\sum_{k^n=1}^{K_i^n} D_{ik}^n}{\sum_{i=1}^{N} K_i^n}\right)\left(\frac{\sum_{i=1}^{N}\sum_{k^n=1}^{K_i^n} E_{ik}^n}{\sum_{i=1}^{N} K_i^n}\right)$$

- λ^e : Average length of a job spell:

$$\overline{J} = \frac{\sum_{i=1}^{N} \frac{\sum_{\ell=1}^{L_i}\left(\sum_{j=1}^{J_{i\ell}-1} D_{i\ell j}\right)}{\sum_{\ell=1}^{L_i} J_{i\ell}-1}}{\sum_{i=1}^{N} 1[\sum_{\ell=1}^{L_i} J_{i\ell} - 1 > 0]}$$

- f_1 :

$$\frac{\sum_{i=1}^{N}\sum_{\ell=1}^{L_i}\sum_{j=1}^{J_{i\ell}} \widetilde{h}_{-i\ell j}\widetilde{S}_{i\ell j}}{\sum_{i=1}^{N}\sum_{\ell=1}^{L_i} J_{i\ell}}$$

  If $\widetilde{h}_{-i\ell j}$ can be defined for all firm-worker combinations this should be the same as the covariance of $\widetilde{S}_{i\ell j}$ and $\widetilde{h}_{-i\ell j}$. This is the case in the simulations. However, notice that for some firm-workers combinations $\widetilde{h}_{-i\ell j}$ cannot be defined in the data. In the data we use the covariance between $\widetilde{h}_{-i\ell j}$ and $\widetilde{S}_{i\ell j}$.

For Wages we have 10 parameters: $E_\theta, \sigma_\theta, \sigma_{xi}, m_1, m_{2,} f_2, f_3, \beta, b_1, b_2$

We use the moments:

- E_θ : Just use

$$\overline{w} \equiv \frac{\sum_{i=1}^{N}\sum_{\ell=1}^{L_i}\sum_{j=1}^{J_{i\ell}}\sum_{t=1}^{T_{i\ell j}} w_{i\ell jt}^{m}}{\sum_{i=1}^{N}\sum_{\ell=1}^{L_i}\sum_{j=1}^{J_{i\ell}} T_{i\ell j}}$$

- $\sigma_\theta^2, \sigma_\xi^2, m1$: We use the decomposition

$$\frac{\sum_{i=1}^{N}\sum_{\ell=1}^{L_i}\sum_{j=1}^{J_{i\ell}}\sum_{t=1}^{T_{i\ell j}} (w_{it}^{m} - \overline{w})^2}{\sum_{i=1}^{N}\sum_{\ell=1}^{L_i}\sum_{j=1}^{J_{i\ell}} T_{i\ell j}} = \frac{\sum_{i=1}^{N}\sum_{\ell=1}^{L_i}\sum_{j=1}^{J_{i\ell}}\sum_{t=1}^{T_{i\ell j}} \left(w_{i\ell jt}^{m} - \overline{w_{i\ell j}}\right)^2}{\sum_{i=1}^{N}\sum_{\ell=1}^{L_i}\sum_{j=1}^{J_{i\ell}} T_{i\ell j}}$$
$$+ \frac{\sum_{i=1}^{N}\sum_{\ell=1}^{L_i}\sum_{j=1}^{J_{i\ell}}\sum_{t=1}^{T_{i\ell j}} (\overline{w_{i\ell j}} - \overline{w_i})^2}{\sum_{i=1}^{N}\sum_{\ell=1}^{L_i}\sum_{j=1}^{J_{i\ell}} T_{i\ell j}} + \frac{\sum_{i=1}^{N}\sum_{\ell=1}^{L_i}\sum_{j=1}^{J_{i\ell}}\sum_{t=1}^{T_{i\ell j}} (\overline{w_i} - \overline{w})^2}{\sum_{i=1}^{N}\sum_{\ell=1}^{L_i}\sum_{j=1}^{J_{i\ell}} T_{i\ell j}}$$

  That is we use each of the three expressions on the right hand side

- f_3 :Cov$(\widetilde{w}_{ilj}, \widetilde{w}_{-ilj})$.

$$\frac{\sum_{i=1}^{N}\sum_{\ell=1}^{L_i}\sum_{j=1}^{J_{i\ell}} \widetilde{w}_{i\ell j}\widetilde{w}_{-i\ell j}}{\sum_{i=1}^{N}\sum_{\ell=1}^{L_i} J_{i\ell}}$$

  If $\widetilde{w}_{-ilj}$ can be defined for all firm-worker combinations this should be the same as the covariance of $\widetilde{w}_{i\ell j}$ and $\widetilde{w}_{-i\ell j}$. This is the case in the simulations. However, notice that for some firm-workers combinations $\widetilde{w}_{-i\ell j}$ cannot be defined in the data. In the data we use the covariance between $\widetilde{w}_{i\ell j}$ and $\widetilde{w}_{-i\ell j}$.

- f_2 :

$$\frac{\sum_{i=1}^{N}\sum_{\ell=1}^{L_i}\sum_{j=1}^{J_{i\ell}} \widetilde{w}_{i\ell j}\widetilde{h}_{-i\ell j}}{\sum_{i=1}^{N}\sum_{\ell=1}^{L_i} J_{i\ell}}$$

  As above. If $\widetilde{h}_{-i\ell j}$ can be defined for all firm-worker combinations this should be the same as the covariance of $\widetilde{w}_{i\ell j}$ and $\widetilde{h}_{-i\ell j}$. This is the case in the simulations. However, notice that for some firm-workers combinations $\widetilde{h}_{-i\ell j}$ cannot be defined in the data. In the data we use the covariance between $\widetilde{w}_{i\ell j}$ and $\widetilde{h}_{-i\ell j}$.

- m_2 :

$$Pr(\overline{w_{i\ell j+1}} < \overline{w_{i\ell j}}) = \frac{\sum_{i=1}^{N}\sum_{\ell=1}^{L_i}\sum_{j=2}^{J_{i\ell}} 1[\overline{w_{i\ell j}} < \overline{w_{i\ell j-1}}]}{\sum_{i=1}^{N}\sum_{\ell=1}^{L_i}(J_{i\ell}-1)}$$

- β, b_1, b_2 : Estimate the following regression using those observations with no left censored tenure (meaning job spells where we do not observe the beginning of the spell)

$$w_{i\ell jt} = \beta_{i\ell j} + \beta_1 E_{i\ell jt} + \beta_2 E^2_{i\ell jt} + \beta_3 TE^2_{i\ell jt} + \epsilon_{i\ell jt}$$

where $\beta_{i\ell j}$ is a job spell fixed effect. We match on the estimates of β_1, β_2, and β_3.

Unobserved heterogeneity In order to identify unobserved heterogeneity we use the following moments.

Variance of employment, job, and non-employment durations taken over individuals

$$\frac{\sum_{i=1}^{N}\left(\frac{\sum_{\ell=1}^{L_i}\left(\sum_{j=1}^{J_{i\ell}} D_{i\ell j}\right)}{L_i} - \overline{L}\right)^2}{\sum_{i=1}^{N} 1[L_i > 0]}$$

$$\frac{\sum_{i=1}^{N}\left(\frac{\sum_{\ell=1}^{L_i}\left(\sum_{j=1}^{J_{i\ell}-1} D_{i\ell j}\right)}{\sum_{\ell=1}^{L_i} J_{i\ell}-1} - \overline{J}\right)^2}{\sum_{i=1}^{N} 1[\sum_{\ell=1}^{L_i} J_{i\ell} - 1 > 0]}$$

$$\frac{\sum_{i=1}^{N}\left(\frac{\sum_{k=1}^{K_i} D^n_{ik}}{K_i} - \overline{K}\right)^2}{\sum_{i=1}^{N} 1[K_i > 0]}$$

Covariance of individual wages and job durations for those jobs that end in a job-to-job transition, where $\widetilde{w} = \frac{1}{N}\sum_{i=1}^{N} \overline{w_i}$

$$\frac{\sum_{i=1}^{N}\left(\frac{\sum_{\ell=1}^{L_i}\left(\sum_{j=1}^{J_{i\ell}-1} D_{i\ell j}\right)}{\sum_{\ell=1}^{L_i} J_{i\ell}-1}\right)(\overline{w_i} - \widetilde{w})}{\sum_{i=1}^{N} 1[\sum_{\ell=1}^{L_i} J_{i\ell} - 1 > 0 \wedge L^m_i > 0]}$$

8 Results

We estimate the model using indirect inference with the auxiliary model described above. Our objective function is the sum of the squared deviation between the simulated model and the data weighted by the inverse of the absolute value of the estimated parameter. We estimate both first using the full sample and then by dividing into four different demographic groups on the basis of gender and education.

The results of this procedure for the model are presented in Tables 3 and 4. Table 3 shows the model fit. One can see that the fit is excellent though we have as many free parameters as we do auxiliary parameters to match. The structural parameters of the model are presented in Table 4. We emphasize that we do not view these as particularly interesting in their own right as they can only be interpreted in the context of the other parameters as many of their values depend upon normalizations we have made in other places. This is particularly important for the λ^e relative to the λ^n parameter as we find a much higher value of the former than the latter. This is due to the fact that non-employment spells tend to be similar in length to employment spells while switching jobs from employment should happen at a lower work since workers with jobs are presumably much pickier. Whether this is due to higher arrival rates on the job or heterogeneity in arrival rates (or reservation values) is hard to identify and we have made a certain normalization through heterogeneity in reservation utility rather than heterogeneity in arrival rates. The primary goal of this project is to explain wages rather than unemployment so we do not view this issue as first order for this paper but worth exploring for other papers that are more concerned with explaining turnover.

We focus on the decomposition of the the amount of total wage variance which is presented in Table 5, where we sequentially eliminate the different sources of wage inequality. First note that prior to the decomposition in the table we get rid of measurement error. The total variance of log wages in the model is 0.125 in the raw data, but it falls to 0.107 after we get rid of the measurement error.

We start with the variance of 0.107 and try to determine which factors contribute to it. Recall that given the issues with the non-employment, eliminating compensating differentials makes less sense than the others. We simulate four different sequences of decompositions (A-D) in which we get rid of alternative sources of inequality. In all cases, our first two steps are the same. In all four cases we eliminate human capital first as it is well known to have little explanatory power (i.e. the R^2 in a Mincer model does not change much when experience and experience squared are dropped) so we view this as well known and less interesting. In addition given the structure of the model human capital largely operates exogenously and separately from other aspects of the model so it does not interact. We show that the small explanatory power is true here as well as human capital explains about 6% of wage variation.

The second experiment we undertake is to rid ourselves of the variation in monopsony powers that firms have over different workers reflected in the bargaining process. Specifically we implement this experiment by setting $\beta = 1$. The reason we do this first is that as discussed above the level of β is is set by normalizations on other parameters (mostly the scale of preferences). This means that changing these other characteristics but holding β fixed makes little sense so we first do this experiment. This gets rid of variation in the model that comes from renegotiation by giving all of the bargaining power to the worker. This lowers the variance of wages by about another 10%. It is important to keep in mind that the fundamental source in the model that leads to this heterogeneity is search frictions-in a perfectly competitive environment firms would have no monopsony power.

We next eliminate the remaining part of search frictions, compensating differentials, and Roy model inequality. For reasons discussed above, eliminating compensating differentials is the most tenuous of these so the most reliable simulations are (A) and (B) in which we eliminate the other two first. To do the simulations

- We eliminate search frictions by allowing people to find the most preferred job immediately.
- We eliminate compensating differentials by setting $v^u_{ij} = \mu^u_j = 0$ for all acceptable jobs (i.e. those that would be taken from nonemployment).
- We eliminate Roy inequality by setting $\sigma_\theta = \sigma_{v^w} = 0$.

The first thing to note is that in all four simulations Roy model inequality is clearly the most important accounting for most of the variation in every case. However, the relative importance of search frictions and compensating differentials varies considerably across the four simulations. That is, this is clearly not an orthogonal decomposition. We do not view this as a weakness of our model but rather as a strength. These different aspects interact in interesting ways. Perhaps most interesting is search frictions. Recall that the monopsony aspect explains 9% of the variation in every case. The remaining amount varies considerably across the specifications. It is about 1% in (A), 5% in (B), 19% in (C) and 3% in (D). What leads to the large differences? The order of the decomposition fundamentally alters the type of match that workers are searching for. In the base case-which corresponds to (B), workers are searching for good matches and care about four different things-the firm specific wage, the firm specific nenpecuniary aspects, the individual$\times$firm type productivity match, and the individual$\times$firm type utility match. When firms are searching for all four of these aspects this leads to 5% variation in earnings. In experiment (A) we first eliminate Roy inequality and then search inequality. Since the we have eliminated the individual$\times$firm type productivity match, which is very important, workers are more interested in searching for the non pecuniary aspects of the match. Perhaps not surprising this aspect is not particularly important for earnings inequality and search only explains 1% of the inequality. In (C) we do the opposite, we first get rid of compensating differentials which means we are getting rid of both the firm specific and ndividual$\times$firm specific nenpecuniary characteristics. In this case workers are searching only for pecuniary aspects of the job and search frictions turn out to be very important-explaining 20% of the variation (30% total if one includes the monopsony part). In (D) we eliminate both the search for non-pecuniary aspects and the individual$\times$firm type productivity match, so that all is being searched for is the firm type productivity. This explains 3% of the variation.

In tables 6 and 7 we present the results for the different demographic groups. In particular we divide into individuals by gender and whether or not they have more than high school or not. First as is clear for 6 the fit is very good for each of the four groups. For the most part the parameters are quite similar across the different groups and the differences make sense. One thing that varies quite a bit that might make less sense is β which is considerably lower for college educated men than it is for the other groups. Taken literally this seems surprising, though as a practical matter what it picks up is that the bargaining process is more important for this group than the others. Intuitively this makes sense (at least to us). It results directly from the fact that the coefficient on tenure squared is larger for college men.

Table 8 presents the decomposition results for the four groups. To save space rather than present the results in four different types of decompositions as in Table 5, we present the results of the different counterfactuals that can be used to produce the decompositions that are shown in Table 5. The main result is robust-Roy model inequality is clearly the most important factor for all four demographic groups. Furthermore the basic results above about search are true as well. The largest drops for search occur if we drop compensating differentials first.

Despite the similarities, there are quite a few factors that are quite different across the different groups. First the level of inequality is very different-the variance is more than twice as large for college men as it is for high school women. Second, the importance of monopsony is quite different-it is a very important factor for college men explaining 20% of the variation while it only explains 10% for high school women. Third, there are large differences in the importance of compensating differentials. Getting rid of it while leaving Roy and Search makes little difference for the variance for men (no difference for college men and a 2% decline for high school men) but leads to large changes for women(13% reduction for college women and a 5% decline for high school women). The relative importance of search frictions depends quite a bit on the experiment we do. For example compensating differentials is relatively important for college women, so relaxing search frictions while keeping tastes for jobs does not have a large effect on inequality. However, if we get rid of compensating differentials first, then the impact of getting rid of search frictions next leads to a large decline for this group (as it does for all of the groups).

9 Conclusions

Not for this draft.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High Wage Workers and High Wage Firms. *Econometrica 67*(2), 251–333.

Bagger, J., F. Fontaine, F. Postel-Vinay, and J.-M. Robin (2011, August). Tenure, experience, human capital and wages: A tractable equilibrium search model of wage dynamics.

Becker, D. (2009). Non-wage job characteristics and the case of the missing margin. Federal Trade Commision.

Bobbio, E. (2010). The Danish Matched Employer-Employee Data. *Unpublished Manuscript*.

Bunzel, H. (2010). The LMDG Data Sets. *Unpublished Manuscript*.

Burdett, K. and D. T. Mortensen (1998). Wage differentials, employer size, and unemployment. *International Economic Review 39*(2), 257–273.

Cahuc, C., F. Postel-Vinay, and J.-M. Robin (2006, March). Wage bargaining with on-the-job search: Theory and evidence. *Econometrica 74*(2), 323–364.

Dey, M. S. and C. J. Flinn (2005, 03). An equilibrium model of health insurance provision and wage determination. *Econometrica 73*(2), 571–627.

Dey, M. S. and C. J. Flinn (2008, January). Household search and health insurance coverage.

Eckstein, Z. and G. Van den Berg (2007). Empirical labor search. *Journal of Econometrics 136*(2), 531–564.

Flinn, C. and J. Heckman (1982). New methods for analyzing structural models of labor force dynamics. *Journal of Econometrics 18*, 115–168.

Gourieroux, C., A. Monfort, and E. Renault (1993, December). Indirect inference. *Journal of Applied Econometrics 8*, S85–S118.

Hall, R. E. and A. B. Krueger (2012, September). Evidence on the incidence of wage posting, wage bargaining, and on-the-job search. *American Economic Journal: Macroeconomics 4*(4), 56–67.

Heckman, J. (1990). Varieties of selection bias. *American Economic Review, Papers and Proceedings 80*.

Heckman, J. and G. Sedlacek (1985). Heterogeneity, aggregation, and market wage functions: An empirical model of self-selection in the labor market. *Journal of Political Economy 93*(6), 1077–1125.

Heckman, J. J. and B. E. Honoré (1990). The empirical content of the roy model. *Econometrica 58*(5), 1121–1149.

Heckman, J. J. and C. R. Taber (2008). Chapter Roy model. Palgrave Macmillan.

Keane, M. P. and K. I. Wolpin (1997). The career decisions of young men. *The Journal of Political Economy 105*(3), 473–522.

Kotlarski, I. (1967). On characterizing the gamma and the normal distributions. *Pacific Journal of Mathematics 20*(1), 69–76.

Postel-Vinay, F. and J.-M. Robin (2002, November). Equilibrium Wage Dispersion with Worker and Employer Heterogeneity. *Econometrica 70*(6), 2295–2350.

Rosen, S. (1987). *Handbook of Labor Economics*, Volume 1, Chapter 12, The Theory of Equalizing Differences,, pp. 641–692.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers 3*(2), 135–146.

Sullivan, P. (2010). A dynamic analysis of educational attainment, occupational choices, and job search. *International Economic Review 51*(1), 289–317.

Sullivan, P. and T. Too (2011, September). Search and non-wage job characteristics.

Villanueva, E. (2007). Estimating compensating wage differentials using voluntary job changes: Evidence from germany. *Industrial and Labor Relations Review 60*(4), 544–561.

Appendix A: Identification

We first consider what can be identified without data on wages. As mentioned in the text we assume that we observe workers from time 0 to T and that all workers begin their working life nonemployed. We observe labor force status during a worker's entire life - that is whether the worker is working and if so, the type of firm for which they work. We only observe wages at a finite number of times. For simplicity assume it is at the integers ($t = 1.0, 2.0, ...$). Let w_{it}^m denote the log of the wage measured at time t if the worker is working at time t.

The easiest parameter to identify is δ which is directly identified from the data as the hazard rate out of employment.

Next consider identification of λ_A^n. Let L_{iA}^n be the length of a nonemployment spell for any worker whose first job was A. For any $\tau_1 < \tau_2 < T$, the conditional density of L_{iA}^n evaluated at τ_1 when we condition on $L_{iA}^n < \tau_2$ is

$$\frac{\lambda_A^n e^{-\lambda_A^n \tau_1}}{1 - e^{-\lambda_A^n \tau_2}}.$$

Since this expression only depends on λ_A^n, it is identified.[12] Following an analogous argument for establishment B, λ_B^n is identified.

The arrival rates of offers on the job can be identified in the same way. That is for a worker who starts at B and moves to A, let L_{iA}^e be their length of employment at job B before taking A. For any $\tau_1 < \tau_2 < T$, the conditional density of L_{iA}^e evaluated at τ_1 when we condition on $L_{iA}^e < \tau_2$ is

$$\frac{\lambda_A^e e^{-\lambda_A^e \tau_1}}{1 - e^{-\lambda_A^e \tau_2}}.$$

One can identify λ_B^e analogously.

In terms of workers choice we can define a "type" by their preferences over jobs. For example one type is:

$$V_{iA0}(\pi_{iA}) < V_{i00} < V_{iB0}(\pi_{iB}), V_{i01} < V_{iA1}(\pi_{iA}) < V_{iB1}(\pi_{iB}).$$

We can define all of these different combinations of types. However, to keep this already complicated section tractable we will assume that there are only 5 different types. Given our panel data, we fully expect the ideas here to extend beyond this case. The five types are:

[12] To see this one can get a closed form solution for λ_A^n by taking the log of the ratio of this density evaluated at two different τ_1 points but the same τ_2 point.

Type	$h=0$	$h=1$
0	$V_{iA0}(\pi_{iA}) < V_{i00}$, $V_{iB0}(\pi_{iB}) < V_{i00}$	NA
B0	$V_{iA0}(\pi_{iA}) < V_{i00} < V_{iB0}(\pi_{iB})$	$V_{iA1}(\pi_{iA}) < V_{i01} < V_{iB1}(\pi_{iB})$
A0	$V_{iB0}(\pi_{iB}) < V_{i00} < V_{iA0}(\pi_{iA})$	$V_{iB1}(\pi_{iB}) < V_{i01} < V_{iA1}(\pi_{iA})$
BA	$V_{i00} < V_{iA0}(\pi_{iA}) < V_{iB0}(\pi_{iB})$	$V_{i01} < V_{iA1}(\pi_{iA}) < V_{iB1}(\pi_{iB})$
AB	$V_{i00} < V_{iB0}(\pi_{iB}) < V_{iA0}(\pi_{iA})$	$V_{i01} < V_{iB1}(\pi_{iB}) < V_{iA1}(\pi_{iA})$

What we have disallowed for is workers changing ordering after human capital accumulates. Again, we see no reason why this would be an issue, but focus on this case for simplicity and it will be true with our empirical specification.

Let D_{it} denote the labor force status of worker i at date t. This can take three different values: A if working for firm A, B if working for firm B, or N for being nonemployed. Now choose any sequence of K time periods, $\tau_1, ..., \tau_K$. Dente the labor market statuses at these time periods $d_{\tau_1}, ..., d_{\tau_K}$. Let *type* index the workers type, then since we know the model for turnover, for each type we can calculate

$$Pr(D_{i\tau_1} = d_{\tau_1}, ..., D_{i\tau_K} = d_{\tau_K} \mid type)$$

from the set of parameters $(\delta, \lambda^n_A, \lambda^n_B, \lambda^e_A, \lambda^e_B)$. Note that this typically will not have a closed form solution, but in principle it is known and in practice it can be approximated arbitrarily well using simulation methods. Moreover since different types make different decisions this function differs across them in nonlinear ways. Then from the data we can identify

$$Pr\left(D_{i\tau_1} = d_{\tau_1}, ..., D_{i\tau_K} = d_{\tau_K}\right) = \sum_{type} Pr(D_{i\tau_1} = d_{\tau_1}, ..., D_{i\tau_K} = d_{\tau_K} \mid type) Pr(type).$$

Note that this is a linear function of $Pr(type)$ where the other components are identified. Thus, we can generate an uncountable number of these equations using different permutations $(d_1, ..., d_T)$ at different time periods. Since these conditional probabilities are nonlinear in type probability, with enough of these linear equations we can identify $Pr(type)$ for all of the different types.

We now incorporate information from wages. We need to define some notation that is cumbersome, but necessary in order for us to be precise. For any worker who is currently working, there are four different states which are relevant for their wages: their current employer, their current human capital, the outside option when their current wage was negotiated, and the level of human capital when the current wage was negotiated. We denote these as functions of the individual and time as $j(i,t), h(i,t), \ell(i,t)$and $h_0(i,t)$, respectively. Then for each integer t at which time the agent is working, we observe

$$w^m_{it} \equiv \log\left(R_{ij(i,t)\ell(i,t)h_0(i,t)}\right) + \log\left(\psi_{h(i,t)}\right) + \xi_{it}$$

where ξ_{it} is i.i.d. measurement error. One can see here the distinction between $h(i,t)$ which is the current level of human capital and $h_0(i,t)$ which is the level of human capital when the current rental rate was negotiated.

To begin we condition on BA types. A random subset of them are easy to identify in the data as they constitute anyone who ever makes a job to job transition from A to B sometime before time T. We condition on this group, but do not make this conditioning explicit in the following for expositional simplicity.

Further condition on BA individuals who

- Are non-employed until time $1-d_1$
- Start working on job A at time $1-d_1$, leave to nonemployment at $1+d_2$
- Are nonemployed until time $2-d_3$ when they start again at a type A firm and they stay through period 2

From this we can identify the joint distribution of

$$(w_{i1}^m, w_{i2}^m)$$

for alternative values of d_1, d_2, and d_3.

Take limits of this conditional distribution as $d_1 \downarrow 0$,$d_2 \downarrow 0$, and $d_3 \downarrow 0$ and we can identify the distribution of

$$\left(\log\left(R_{iA00}\right)+\xi_{i1}, \log\left(R_{iA00}\right)+\xi_{i2}\right)$$

for the BA types. Both of these wages will be $\log\left(R_{iA00}\right)$ because the people have not had enough time to accumulate human capital or get an outside offer. Notice, that since $\psi_0=1$ then $R_{iA00}=W_{iA00}$, i.e. the rental rate is equal to the wage paid. Using Kotlarski's lemma (Kotlarski 1967) we can identify the the marginal distributions of both the measurement error and $\log\left(R_{iA00}\right)$.

Consider the same group and continue to take $\delta_1 \downarrow 0$ and $\delta_3 \downarrow 0$, but allow d_2 to vary then let $\phi_\xi(s)$ be the characteristic function of the measurement error which is identified. Assuming that the characteristic function of the measurement error does not vanish, we can identify

$$\frac{E\left(e^{\imath s w_{i2}^m}\right)}{\phi_\xi(s)} \approx e^{-\lambda_h \delta_2} E\left(e^{\imath s \log(R_{iA00})}\right)+\left(1-e^{-\lambda_h d_2}\right) E\left(e^{\imath s \log(R_{iA01})}\right).$$

By varying δ_2 we can identify λ_h and $E\left(e^{\imath slog(R_{iA01})}\right)$.[13]

Now consider individuals whose first job is A, they stay at the job through $t = 1$ and $t = 2$, but eventually move to B. For them at each time period $t = 1$ and $t = 2$ there are 4 potential labor statuses:

1. Never got outside offer, $h = 0$
2. Outside offer from A, $h = 0$
3. Never got outside offer, $h = 1$
4. Outside offer from A, $h = 1$

They can not have gotten an offer from firm B or they would have accepted it. When a worker at A gets an outside offer from another A firm, the bargained wage satisfies

$$V_{iAh}(R_{iAAh}) = \beta V_{iAh}(\pi_{iA}) + (1-\beta)V_{iAh}(\pi_{iA}),$$

so $R_{iAAh} = \pi_{iA}$. Thus for status 4, the log wage will be $\pi_{iA} + h_1$ regardless of whether the outside offer or human capital innovation came first. Let $p_{r_1 r_2}(d_1)$ represent the probability that the status is r_1 in the first period and r_2 in the second for someone who started the job at time $1 - d_1$. Then let $\phi_{w^m_{i1},w^m_{i2}}(s_1, s_2; d_1)$ be the characteristic function of the joint distribution of (w^m_{i1}, w^m_{i2}) conditional on BA people who started job A at time $1 - d_1$ and

[13] To see how, take the ratio of the derivatives of this function in terms of δ_2 at two different values of δ_2 and it will be a known function of λ_h.

The derivative of that expression with respect to δ_2 is

$$-\lambda_h e^{-\lambda_h d_2} E\left(e^{\imath slog(R_{iA00})}\right) + \lambda_h e^{-\lambda_h d_2} E\left(e^{\imath slog(R_{iA01})}\right) = \lambda_h e^{-\lambda_h d_2}\left[E\left(e^{\imath slog(R_{iA01})}\right) - E\left(e^{\imath slog(R_{iA00})}\right)\right]$$

Now take the ratio of this at two different values of δ_2 say δ_2^a and δ_2^b then

$$\begin{aligned}\Delta(\delta_2^a, \delta_2^b) \equiv & \frac{\lambda_h e^{-\lambda_h d_2^a}\left[E\left(e^{\imath slog(R_{iA01})}\right) - E\left(e^{\imath slog(R_{iA00})}\right)\right]}{\lambda_h e^{-\lambda_h d_2^b}\left[E\left(e^{\imath slog(R_{iA01})}\right) - E\left(e^{\imath slog(R_{iA00})}\right)\right]} \\ = & e^{\lambda_h\left(\delta_2^b - \delta_2^a\right)}\end{aligned}$$

$\Delta(\delta_2^a, \delta_2^b)$ is directly identified from the data and

$$\lambda_h = \frac{\log\left(\Delta(d_2^a, d_2^b)\right)}{\delta_2^b - \delta_2^a}$$

still employed there through time $t = 2$. Then we can identify

$$
\begin{aligned}
&\frac{\phi_{w_{i1}^m, w_{i2}^m}(s_1, s_2; d_1)}{\phi_\varepsilon(s_1)\phi_\varepsilon(s_2)} \\
=&p_{11}\left(d_1\right) E\left(e^{\imath(s_1+s_2)\log(R_{iA00})}\right) + p_{12}\left(d_1\right) E\left(e^{\imath(s_1\log(R_{iA00})+s_2\log(\pi_{iA}))}\right) \\
&+ p_{13}\left(d_1\right) E\left(e^{\imath[(s_1+s_2)\log(R_{iA00})+s_2\log(\psi_1)]}\right) + p_{14}\left(d_1\right) E\left(e^{\imath(s_1\log(R_{iA00})+s_2\log(\pi_{iA}\psi_1))}\right) \\
&+ p_{22}\left(d_1\right) E\left(e^{\imath(s_1+s_2)\log(\pi_{iA})}\right) + p_{24}\left(d_1\right) E\left(e^{\imath[(s_1+s_2)\pi_{iA}+s_2\log(\psi_1)]}\right) \\
&+ p_{33}(d_1)E\left(e^{\imath(s_1+s_2)(\log(R_{iA00})+\log(\psi_1))}\right) + p_{34}(d_1)E\left(e^{\imath[s_1(\log(R_{iA00})+log(\psi_1))+s_2(log(\pi_{iA})+log(\psi_1))]}\right) \\
&+ p_{44}(d_1)E\left(e^{\imath(s_1+s_2)(log(\pi_{iA})+log(\psi_1))}\right)
\end{aligned}
$$

We have shown that the transition parameters are identified from which we can identify $p_{j\ell}(d_1)$. We can vary d_1 continuously. Given that the form of $p_{j\ell}(d_1)$ is nonlinear that we can solve for each of the expected values so by varying it we can get a system of linear equations and invert it to obtain each of the expected values in the equation above. From this it is clear that ψ_1 is identified as well as the characteristic function (and thus joint distribution) of $\log\left(R_{iA00}\right)$ and π_{iA}.

Using similar logic and 8 periods of wage data we can identify the joint distribution of

$$R_{iA00}, \pi_{iA}, R_{iBA0}, R_{iB00}, \pi_{iB}, R_{iA01}, R_{iBA1}, R_{iB01}).$$

Given the number of different combinations of labor market statuses we do not write this out completely but the form will be similar to the example above.

We can use the exact same strategy to identify

$$(R_{iA00}, \pi_{iA}, R_{iAB0}, R_{iB00}, \pi_{iB}, R_{iA01}, R_{iAB1}, R_{iB01})$$

for the AB group.

Next consider the $A0$ type. This is more complicated as we can not directly distinguish an $A0$ type from an AB or BA type that never worked for a type B firm during our sample period. For the $A0$ type, there are only three relevant wages R_{iA00}, R_{iA01}, and π_{iA}. To see how to identify this joint distribution consider workers who begin their first job at A at $1-\delta_1$, end that job at $1+\delta_2$, begin a new job at A at $2-\delta_3$ and are still working at that job at time 3. Send $\delta_1 \downarrow 0$ and $\delta_3 \downarrow 0$ as above. We know that at the first period the worker has no human capital and no outside offers. At period 2 they have no outside offer, but may have human capital. At period three they can be in any of the four states listed above. Being someone loose with notation let P_i be a labor market profile for the three periods and let $p(\delta_2)$ be the particular profile we are conditioning on (i.e. being unemployed until 1-δ_1, working at A from $1-\delta_1$ until $1+\delta_2$, being unemployed until $1-\delta_3$, then working at job A through period 3). Then we can identify

$$\begin{aligned}
&\phi_{w^m_{i1},w^m_{i2},w^m_{i3}}(s_1,s_2,s_3;\delta_2)\\
=&Pr(type=A0 \mid P_i=p(\delta_2))E\left(e^{\imath\left(s_1w^m_{i1}+s_2w^m_{i2}+s_3w^m_{i3}\right)} \mid type=A0, P_i=p(\delta_2)\right)\\
&+Pr(type=AB \mid P_i=p(\delta_2))E\left(e^{\imath\left(s_1w^m_{i1}+s_2w^m_{i2}+s_3w^m_{i3}\right)} \mid type=AB, P_i=p(\delta_2)\right)\\
&+Pr(type=BA \mid P_i=p(\delta_2))E\left(e^{\imath\left(s_1w^m_{i1}+s_2w^m_{i2}+s_3w^m_{i3}\right)} \mid type=BA, P_i=p(\delta_2)\right).
\end{aligned}$$

We have previously shown that everything in the above expression other than the conditional characteristic function $E\left(e^{\imath\left(s_1w^m_{i1}+s_2w^m_{i2}+s_3w^m_{i3}\right)} \mid type=A0, P_i=p\right)$ is identified, thus it is identified from this expression.

We now continue to condition on the case above with $\delta_1 \approx \delta_3 \approx 0$ and let $p_{r_2r_3}(\delta_1)$ represent the probability that the worker has labor market status r_2 in the second period and status r_3 in period 3 (statuses are listed above). From this we can identify (where we condition on type=A0 but do not make it explicit)

$$\begin{aligned}
&\frac{E\left(e^{\imath\left(s_1w^m_{i1}+s_2w^m_{i2}+s_3w^m_{i3}\right)} \mid type=A0, P_i=p\right)}{\phi_\varepsilon(s_1)\phi_\varepsilon(s_2)\phi_\varepsilon(s_3)}\\
=&p_{11}\left(\delta_2\right)E\left(e^{\imath(s_1+s_2+s_3)\log(R_{iA00})}\right)+p_{12}\left(\delta_2\right)E\left(e^{\imath((s_1+s_2)\log(R_{iA00})+s_3\log(\pi_{iA}))}\right)\\
&+p_{13}\left(\delta_2\right)E\left(e^{\imath[(s_1+s_2+s_3)\log(R_{iA00})+s_3\log(\psi_1)]}\right)+p_{14}\left(\delta_2\right)E\left(e^{\imath((s_1+s_2)\log(R_{iA00})+s_3(log(\pi_{iA})+log(\psi_1)))}\right)\\
&+p_{33}\left(\delta_2\right)E\left(e^{\imath(s_1\log(R_{iA00})+(s_2+s_3)\log(R_{iA01}))}\right)+p_{34}\left(\delta_2\right)E\left(e^{\imath(s_1\log(R_{iA00})+s_2\log(R_{iA01})+s_3\log(\pi_{iA}\psi_1))}\right).
\end{aligned}$$

Using the same logic as above we can identify all of the expected value terms on the right hand side including $E\left(e^{\imath(s_1\log(R_{iA00})+s_2\log(R_{iA01})+s_3(log(\pi_{iA})+log(\psi_1)))}\right)$. From this we know the joint distribution of $(R_{iA00},R_{iA01},\pi_{iA})$ for the A0 types.

An analogous argument shows identification of $(R_{iB00},R_{iB01},\pi_{iB})$ for the B0 types.

We have shown that we can identify wages and revealed preference about job choices. The next question is whether we can identify the individual utility functions and in particular the bargaining parameter β. The answer turns out to be no. To see why, suppose that we could observe all of the wages for a particular BA worker (which is similar to observing the distribution of wages after filtering out the measurement error). We could observe the 8 wages $(W_{iA00},\pi_{iA},W_{iBA0},W_{iB00},\pi_{iB},W_{iA01},W_{iBA1},W_{iB01})$. Assume even further that the utility function takes the form

$$U_{ij}(W) = \log(w) + v_{ij}.$$

Even in this restrictive case, β is not identified. Writing down the wage equations for the 6 wages we can get for human capital level 1:

$$\begin{aligned}
(\rho+\delta)\,V_{iB1}(\pi_{iB}) =& \log(\pi_{iB}) + \log(\psi_1) + v_{iB} + \delta V_{i01} \\
(\rho+\delta+\lambda_B^e)\,V_{iA1}(\pi_{iA}) =& \log(\pi_{iA}) + \log(\psi_1) + v_{iA} \\
&+ \lambda_B^e\left[\beta V_{iB1}(\pi_{iB}) + (1-\beta)\,V_{iA1}(\pi_{iA})\right] + \delta V_{i01} \\
(\rho+\lambda_A^n+\lambda_B^n)\,V_{i01} =& u_{i01} + \lambda_A^n\left[\beta V_{iA1}(\pi_{iA}) + (1-\beta)\,V_{i01}\right] \\
&+ \lambda_B^n\left[\beta V_{iB1}(\pi_{iB}) + (1-\beta)\,V_{i01}\right] \\
(\rho+\delta+\lambda_B^e+\lambda_A^e)&\left[\beta V_{iA1}(\pi_{iA}) + (1-\beta)\,V_{i01}\right] \\
=& \log(w_{iA01}) + v_{iA} + \lambda_A^e V_{iA1}(\pi_{iA}) \\
&+ \lambda_B^e\left[\beta V_{iB1}(\pi_{iB}) + (1-\beta)\,V_{iA1}(\pi_{iA})\right] + \delta V_{i01} \\
(\rho+\delta+\lambda_B^e+\lambda_A^e)&\left[\beta V_{iB1}(\pi_{iB}) + (1-\beta)\,V_{i01}\right] \\
=& \log(w_{iB01}) + v_{iB} + \lambda_A^e\left[\beta V_{iB1}(\pi_{iB}) + (1-\beta)\,V_{iA1}(\pi_{iA})\right] \\
&+ \lambda_B^e V_{iB1}(\pi_{iB}) + \delta V_{i01} \\
(\rho+\delta+\lambda_B^e)&\left[\beta V_{iB1}(\pi_{iB}) + (1-\beta)\,V_{iA1}(\pi_{iA})\right] \\
=& \log(w_{iBA1}) + v_{iB} + \lambda_B^e V_{iB1}(\pi_{iB}) + \delta V_{i01}
\end{aligned}$$

For human capital level 0

$$
\begin{aligned}
(\rho+\delta+\lambda_h)\,V_{iB0}(\pi_{iB}) =& \log(\pi_{iB}) + v_{iB} + \delta V_{i0} + \lambda_h V_{iB1}(\pi_{iB}) \\
(\rho+\delta+\lambda_B^e+\lambda_h)\,V_{iA0}(\pi_{iA}) =& \log(\pi_{iA}) + v_{iA} + \lambda_B^e\left[\beta V_{iB0}(\pi_{iB}) + (1-\beta)\,V_{iA0}(\pi_{iA})\right] \\
& + \delta V_{i0} + \lambda_h V_{iA1}(\pi_{iA}) \\
(\rho+\lambda_A^n+\lambda_B^n)\,V_{i00} =& u_{i00} + \lambda_A^n\left[\beta V_{iA0}(\pi_{iA}) + (1-\beta)\,V_{i00}\right] \\
& + \lambda_B^n\left[\beta V_{iB0}(\pi_{iB}) + (1-\beta)\,V_{i00}\right] \\
(\rho+\delta+\lambda_B^e+\lambda_A^e+\lambda_h)&\left[\beta V_{iA0}(\pi_{iA}) + (1-\beta)\,V_{i00}\right] \\
=& \log(w_{iA00}) + v_{iA} + \lambda_A^e V_{iA0}(\pi_{iA}) \\
& + \lambda_B^e\left[\beta V_{iB0}(\pi_{iB}) + (1-\beta)\,V_{iA0}(\pi_{iA})\right] + \delta V_{i00} + \lambda_h V_{iA1}(w_{iA00}) \\
& + \lambda_h \frac{\log(w_{iA00}) + \log(\psi_1) + v_{iA} + \delta V_{i01}}{\rho+\delta+\lambda_B^e+\lambda_A^e} \\
& + \lambda_h \frac{\lambda_A^e V_{iA1}(\pi_{iA}) + \lambda_B^e\left[\beta V_{iB1}(\pi_{iB}) + (1-\beta)\,V_{iA0}(\pi_{iA})\right]}{\rho+\delta+\lambda_B^e+\lambda_A^e} \\
(\rho+\delta+\lambda_B^e+\lambda_A^e+\lambda_h)&\left[\beta V_{iB0}(\pi_{iB}) + (1-\beta)\,V_{i00}\right] \\
=& \log(w_{iB00}) + v_{iB} + \lambda_A^e\left[\beta V_{iB0}(\pi_{iB}) + (1-\beta)\,V_{iA0}(\pi_{iA})\right] \\
& + \lambda_B^e V_{iB0}(\pi_{iB}) + \delta V_{i00} + \lambda_h V_{iB1}(w_{iB00}) \\
& + \lambda_h \frac{\log(w_{iB00}) + \log(\psi_1) + v_{iB} + \delta V_{i01}}{\rho+\delta+\lambda_B^e+\lambda_A^e} \\
& + \lambda_h \frac{\lambda_A^e\left[\beta V_{iB0}(\pi_{iB}) + (1-\beta)\,V_{iA0}(\pi_{iA})\right] + \lambda_B^e V_{iB0}(\pi_{iB})}{\rho+\delta+\lambda_B^e+\lambda_A^e} \\
(\rho+\delta+\lambda_B^e+\lambda_h)&\left[\beta V_{iB0}(\pi_{iB}) + (1-\beta)\,V_{iA0}(\pi_{iA})\right] \\
=& \log(w_{iBA0}) + v_{iB} + \lambda_B^e V_{iB0}(\pi_{iB}) + \delta V_{i00} \\
& \lambda_h \frac{\log(w_{iBA0}) + \log(\psi_1) + v_{iB} + \lambda_B^e V_{iB1}(\pi_{iB}) + \delta V_{i01}}{\rho+\delta+\lambda_B^e}
\end{aligned}
$$

This is 12 equations in 11 unknowns

$$
V_{i00}, V_{i01}, V_{iB0}(\pi_{iB}), V_{iB1}(\pi_{iB}), V_{iA0}(\pi_{iA}), V_{iA1}(\pi_{iA}), u_{i0}, u_{i1}, v_{iA}, v_{iB},
$$

and β. Beyond this we can clearly get a normalization on flow utility, say normalizing u_{i00} to zero, so this is 12 equations in 10 unknowns. However, three of the equations are linearly dependent so we actually have only 9 separate equations and 10 unknowns. Specifically, it is straightforward to show that

$$\begin{aligned}
\log\left(w_{iBA1}\right)-\left(\log\left(w_{iB01}\right)-\log\left(w_{iA01}\right)\right) &= \log\left(\pi_{iA}\right)+\log\left(\psi_{1}\right) \\
\log\left(w_{iBA0}\right)-\left(\log\left(w_{iB00}\right)-\log\left(w_{iA00}\right)\right) &= \log\left(\pi_{iA}\right) \\
\log\left(w_{iBA0}\right) &= \log\left(w_{iBA1}\right)-\log(\psi_{1})
\end{aligned}$$

Thus, the model is not identified and in particular as long as $\beta < 1$ we could choose it to any value andsolve all equations. However $\beta = 1$ can be tested as it would imply, for example, $w_{iA00} = \pi_{iA}$.

Appendix B: Data

Data Selection We define labor market entry to be the month of graduation from the highest completed education recorded.[14] We disregard spells that are before this date. If the worker after the date of highest completed education is observed in education the worker is disregarded. E.g. if the highest recorded education for a worker is high school and he graduated in 2001 and we later observe him in education, say in 2003 then we delete him. Workers with changing codes for highest completed education and where age minus education length is less than 5 years are also disregarded. We censor workers after age 55.

Temporary non-employment (unemployment and non-participation) spells shorter than 13 weeks where the previous and next establishment id are the same as one employment spell, i.e. unemployment and non-participation spells are treated as one type of spells. Short unemployment or non-participation spells between two employment spells shorter than 3 weeks are allocated to the last of the two employment spells.

We censor workers when they enter a self-employment state. We delete workers that have gaps in their spell histories. This could arise if the worker for some reason have missing IDA data in a given year. Wages are detrended in logs (but so far not trimmed). We label the states unemployment, retirement and non-participation as non-employment. For some of the employed workers in the final sample we do not observe the establishment ID. However, this is a relatively small fraction, see Table 2. In the calculations of the moments we will take this into account, and only use observations for which we do observe the establishment.

[14] We have information on highest completed education back to 1969, so highest completed education is missing for workers who took it before 1969. Also, immigrants and workers who never finished primary school have missing values. We keep this workers since we suspect that the problems with immigrants and workers who never finished primary school are quit small, and workers who took there education before 1969 have entered the labor market.

E.g. if a transition happens from establishment 1 to an unknown establishment we will not count this as a transition. Likewise, if a transition from an unknown establishment to firm 1 will not be counted as a transition.

In the identification strategy we heavily rely on the fact that observed job to job transitions are actually voluntary. One might suspect that workers in closing establishments might move to a new establishment without actually preferring it compared to the old one. In order to avoid drawing inference from such observations we do not count job to job transitions from an establishment in the year that it closes. E.g. if a workers is employed in establishment 1 in week 1 to 40 in 1995 (and we do not observe the establishment in the data after 1995) and in week 41 he is employment in establishment 2 then we do not count that as job to job separation for establishment 1 nor a job to job hire in establishment 2, although we will count them as a separation and a hire. However, if the worker had transitioned into non-employment we would have counted a job destruction. This limits the sample to 2002, since we cannot observe if 2003 was the last year of which an establishment was alive. This gives us the final sample.

Estimating Labor Market Entry: We observe graduation times from 1971 and forward. Our sample starts in 1985 which means that we observe some individuals with labor market entry before 1971 (around 1/3 of all workers). We therefore need to approximate the entry year. We use population data (not just those in our sample) graduating in 1971 and 1972 to derive the age distribution at graduation time by gender-education group. This gives us around 70 groups. However a few workers in the sample cannot be matched, so we use a more rough groups for those. We now use the gender-education specific graduation age distribution conditional on the fact that we know the individual did not graduate after 1970. If the minimum age in the estimated distribution implies entry after 1970 we set entry to 1970. This is the case for 1 percent of the workers we approximate.

Estimating Experience Experience is observed yearly from 1964 (although with different degrees of precision). We define experience to be experience accumulated from labor market entry. Given that we have approximated entry time we need to approximate experience up to 1970. To do this we again use those entering the labor market in 1971 and 1972. There are several ways to do this. However, one of the simplest is to calculate the yearly mean experience increase. Assuming that individuals either work full time or not at all we approximate experience up to 1970 using a binomial distribution with probability estimated by gender-education-time since entry groups. We thus divide workers into 4 groups based on time since entry. These are 1-5, 6-10, 11-15, and above 15 years since entry. An example of a group could be female Kindergarten teachers with 1 to 5 years in the labor market.

Table 1: Overview of data creation

	Number of workers	Number of Firms	Number of Spells	Number of Establishments	Fraction Employed
Merged data	5,116,625	455,054	60,914,366	610,130	
Censoring at age 56	4,348,157	446,957	56,090,810	596,777	
Delete workers with all missing educational variables (hffsp)	4,147,463	445,298	54,714,582	594,070	
Delete workers with gaps in educational variables (hffsp)	4,147,463	445,298	54,714,582	594,070	
Delete obs. below age 19	3,967,419	420,223	51,494,700	550,388	
Labor market entry	3,967,419	400,670	43,406,322	517,710	
Delete under education	3,461,332	386,894	37,007,709	496,792	
Changing hffsp codes	3,346,165	385,833	36,778,315	495,453	
To early labor market entry	3,322,994	385,289	36,540,726	494,727	
Clean for temporary non-employment	3,313,246	384,988	29,495,476	494,275	
Censoring self-employment	3,118,361	368,839	26,920,923	472,219	
Censoring retirement	3,041,336	368,747	26,462,593	472,095	
Delete workers with gabs	2,969,454	365,959	25,630,532	467,983	
Delete workers with missing experience	2,969,454	365,959	25,630,532	467,983	
Average over yearly cross-sections					
Merged data	3,717,692	129,506	3,717,692	167,882	0.60
Censoring at age 56	2,955,775	124,698	2,955,775	161,952	0.68
Delete workers with all missing educational variables (hffsp)	2,894,716	124,084	2,894,716	161,285	0.69
Delete workers with gaps in educational variables (hffsp)	2,894,716	124,084	2,894,716	161,285	0.69
Delete obs. below age 19	2,700,370	116,688	2,700,370	153,681	0.70
Labor market entry	2,375,046	110,346	2,375,046	146,898	0.71
Delete under education	2,143,752	106,593	2,143,752	142,682	0.72
Changing hffsp codes	2,137,499	106,470	2,137,499	142,548	0.72
To early labor market entry	2,127,965	106,353	2,127,965	142,423	0.72
Clean for temporary non-employment	2,120,247	106,635	2,120,247	142,777	0.72
Censoring self-employment	1,872,217	100,595	1,872,217	136,301	0.78
Censoring retirement	1,763,850	100,568	1,763,850	136,269	0.83
Delete workers with missing experience	1,722,319	99,802	1,722,319	135,375	0.83

Table 2: Summery Statistics: Pooled Cross-sections

	Mean	Std. Dev.	Min.	Max.
Number of years in sample	11.09	6.27	1	19
Number of Establishments per worker	2.74	1.87	1	18
Female	0.49			
Average years of education	11.66	3.18	7	20
Age	38.33	9.63	19	55
Employed	0.84			
Public Employed	0.32			
Missing Establishment ID	0.005			
Experience	13.18	9.22	0.00	40.92
Log Wages	4.54	0.34	0.04	9.07

Table 3: Auxiliary Model and Estimates: Full Sample

Moment	Data	Model (Immed Possible)
Avg. Length Emp. Spell	377	382
Avg. Length Nonnemp. Spell	91.4	91.6
Avg. Length Job	108	107
$E(\widetilde{S}_{i\ell j}\widetilde{h}_{-i\ell j})$	0.021	0.022
Between Persons	0.0808	0.0807
Between Jobs	0.0289	0.0290
Within Job	0.0151	0.0151
Sample mean w_{it}	4.51	4.51
$\mathrm{E}(\widetilde{w}_{it}\widetilde{w}_{-it})$	0.00393	0.00393
$E(\widetilde{w}_{it}\widetilde{h}_{-it})$	0.00152	0.00152
Fraction Wage Drops	0.400	0.395
Coeff Exper	0.0250	0.0252
Coeff Exper^2	-0.00031	-0.00031
Coef Tenure^2	-0.00047	-0.00047
Var(Nonemployment)	16000	15988
Cov($\overline{w}_i$, Non-employment)	-3.53	-3.53
Var(Employment Dur)	102000	101067
Invol Job to Job	0.200	0.200

Table 4: Parameter Estimates: Full Sample

Parameter	
d_0	-2.94
λ^n	0.97
λ^e	2.10
E_θ	4.35
σ_θ	0.231
σ_ξ	0.134
f_1	4.78
f_2	0.141
f_3	-0.055
σ_{v^w}	0.188
α	4.36
b_1	0.018
b_2	-0.0008
β	0.825
P^*	0.421
σ_{ν_0}	0.472
γ_θ	-0.672
σ_d	2.07

Table 5: Model Decompositions: Variance of log wages, Full Sample

(A)		(B)		(C)		(D)	
Total	0.107	Total	0.107	Total	0.107	Total	0.107
No HC	0.101	No HC	0.101	No HC	0.101	No HC	0.101
No Monop	0.091	No Monop	0.091	No Monop	0.091	No Monop	0.091
No Roy	0.008	No Search	0.086	No Comp	0.089	No Comp	0.089
No Search	0.007	No Roy	0.007	No Search	0.069	No Roy	0.003

Table 6: Auxliary Model and Estimates: Demographic Groups

	Col Men		HS Men		Col Women		HS Women	
Moment	Data	Model	Data	Model	Data	Model	Data	Model
Avg. Length Emp. Spell	430	427	382	377	392	391	346	349
Avg. Length Nonnemp. Spell	60.1	60.0	80	80	63	63	119	119
Avg. Length Job	120	120	103	104	109	109	109	108
$E(\widetilde{S}_{i\ell j}\widetilde{h}_{-i\ell j})$	0.0258	0.0258	0.0225	0.0216	0.0186	0.0186	0.0218	0.0230
Between Persons	0.0983	0.0982	0.0620	0.0556	0.0505	0.0503	0.0452	0.0449
Between Jobs	0.0325	0.0329	0.0313	0.0308	0.0258	0.0259	0.0258	0.0260
Within Job	0.0186	0.0186	0.0142	0.0142	0.0145	0.0145	0.0151	0.0151
Sample mean w_{it}	4.78	4.78	4.57	4.57	4.49	4.49	4.33	4.33
E$(\widetilde{w}_{it}\widetilde{w}_{-it})$	0.0056	0.0056	0.0056	0.0056	0.0029	0.0029	0.0027	0.0027
$E(\widetilde{w}_{it}\widetilde{h}_{-it})$	0.0026	0.0026	0.0019	0.0019	0.0014	0.0014	0.0012	0.0012
Fraction Wage Drops	0.335	0.331	0.408	0.401	0.405	0.418	0.413	0.417
Coeff Exper	0.0427	0.0413	0.0251	0.0248	0.025	0.025	0.0189	0.0186
Coeff Exper^2	-0.00067	-0.00069	-0.00028	-0.00028	-0.00036	-0.00036	-0.00022	-0.00022
Coef Tenure^2	-0.00076	-0.00076	-0.00050	-0.00050	-0.00046	-0.00046	-0.00031	-0.00031
Var(Nonemployment)	7830	7862	12000	11978	8513	8550	22800	22718
Cov($\overline{w}_i$, Non-employment)	-2.04	-2.04	-2.52	-2.51	-1.46	-1.47	-1.52	-1.52
Var(Employment Dur)	102000	98957	107000	110084	98100	101560	94700	96960
Invol Job to Job	0.182	0.182	0.243	0.242	0.086	0.086	0.218	0.219

Table 7: Parameter Estimates By Group

Parameter	Col Men	HS Men	Col Women	HS Women
d_0	-4.47	-2.91	-3.91	-2.56
λ^n	1.84	1.17	1.82	0.781
λ^e	2.28	2.41	2.43	1.678
E_θ	4.60	4.49	4.42	4.23
σ_θ	0.255	0.178	0.141	0.124
σ_ξ	0.121	0.125	0.125	0.140
f_1	11.18	9.51	1.38	5.62
f_2	0.186	0.161	0.147	0.120
f_3	-0.0240	-0.0205	0.169	0.0168
σ_{v^w}	0.185	0.175	0.198	0.193
α	13.98	3.74	1.798	3.39
b_1	0.0412	0.0078	0.014	-0.003
b_2	-0.0034	-0.000	-0.0015	0.001
β	0.564	0.875	0.758	0.905
P^*	0.586	0.418	0.226	0.387
σ_{ν_0}	0.370	0.855	0.486	0.760
γ_θ	4.60	4.49	4.42	4.23
σ_d	2.59	2.59	2.59	2.02

Table 8: Model Decompositions: Variance of log wages by Demographic Groups

	Col Men	HS Men	Col Women	HS Women
Total	0.135	0.092	0.075	0.066
No HC	0.128	0.088	0.074	0.062
No Monop	0.101	0.070	0.063	0.056
No Roy	0.012	0.010	0.008	0.006
No Search	0.085	0.064	0.061	0.048
No Comp	0.101	0.068	0.053	0.053
No Roy/Search	0.008	0.009	0.007	0.005
No Roy/Comp	0.005	0.004	0.004	0.003
No Search/Comp	0.080	0.047	0.034	0.029

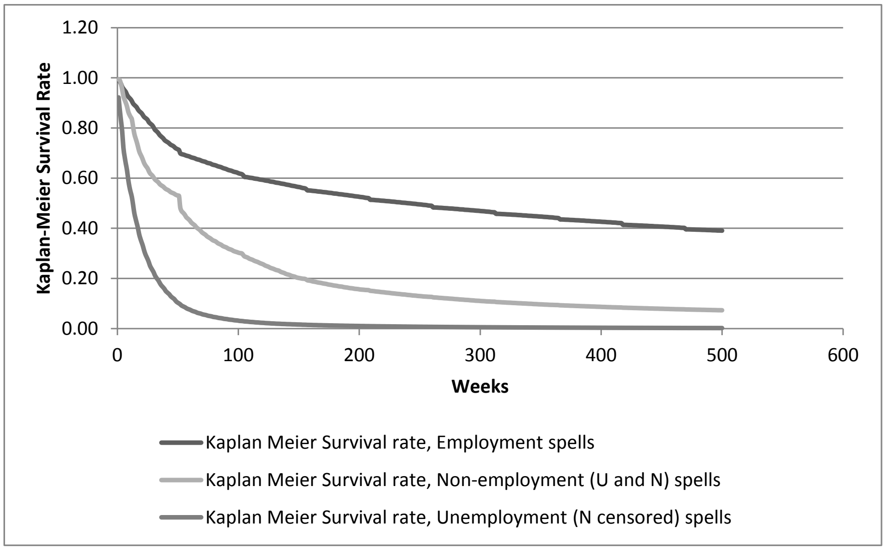

Figure 1: Survival Plots of Employment, Unemployment, and Non-Employment Spells