Appedices D and E: Estimation of a Roy/Search/Compensating Differential Model of the Labor Market

Christopher Taber

University of Wisconsin-Madison

Rune Vejlin

Aarhus University and IZA

March 7, 2020

Appendix D Identification

In this section we provide the details regarding section 4 in the paper. We begin with Lemma D.1 in section D.1. The proof of theorem 1 is presented in D.2. Section D.3 gives the proof for Theorem 2. To give intuition it begins with the proof for a much simplified version of the model. D.4 gives the proof of Theorem 3. Sections D.5 and D.6 respectively present Theorems D.1 and D.2 and prove them. These theorems are the analogues of Theorems 1 and 2 but modify the model and data by allowing heterogeneity in δ_i but require the econometrician to have an infinitely long time span to view the data. These proofs are very similar to the proofs of Theorems 1 and 2. The proof of Theorem 3 goes through exactly with the assumptions of Theorem D.2 so we do not explicitly show that.

D.1 Lemma D.1

Suppose that the hazard function takes the form

$$h\left(\tau\right) = aw\left(\tau\right) + b\left[1 - w\left(\tau\right)\right],$$

where

$$w\left(\tau\right) = \frac{1}{1 + ce^{-\lambda\tau}},$$

and $\lambda \neq 0$ and $c > 0$ then a, b, c, and λ are all identified from the hazard function.

Proof of Lemma D.1

First, note that

$$\frac{1 - w\left(\tau\right)}{w\left(\tau\right)} = ce^{-\lambda\tau}.$$

We first show that this relationship and the hazard function equation identifies a and b. Suppose not, suppose there exists another $a^* \neq a$ and/or $b^* \neq b$ that has this same property. In that case there is a function $w(\tau; a^*, b^*)$ for which

$$h(\tau) = a^* w(\tau; a^*, b^*)) + b^* \left(1 - w(\tau; a^*, b^*)\right),$$

and

$$\frac{1 - w(\tau; a^*, b^*)}{w(\tau; a^*, b^*)} = c^* e^{-\lambda^* \tau},$$

for some $\lambda^*$and c^*.

Since

$$w(\tau; a^*, b^*) = \frac{h(\tau) - b^*}{a^* - b^*},$$

then

$$\begin{aligned}\frac{1 - w(\tau; a^*, b^*)}{w(\tau; a^*, b^*)} &= \frac{1 - \frac{h(\tau) - b^*}{a^* - b^*}}{\frac{h(\tau) - b^*}{a^* - b^*}} \\ &= \frac{a^* - h(\tau)}{h(\tau) - b^*} \\ &= \frac{a^* - a\frac{1}{1+ce^{-\lambda\tau}} - b\frac{ce^{-\lambda\tau}}{1+ce^{-\lambda\tau}}}{a\frac{1}{1+ce^{-\lambda\tau}} + b\frac{ce^{-\lambda\tau}}{1+ce^{-\lambda\tau}} - b^*} \\ &= \frac{a^*\left(1 + ce^{-\lambda\tau}\right) - a - bce^{-\lambda\tau}}{a + bce^{-\lambda\tau} - b^*\left(1 + ce^{-\lambda\tau}\right)} \\ &= \frac{(a^* - a) + (a^* - b)\, ce^{-\lambda\tau}}{(a - b^*) + (b - b^*)\, ce^{-\lambda\tau}},\end{aligned}$$

does not take the functional form above unless $b = b^*$ and $a = a^*$.[1] To derive this expression we first plug in the derivation of $w(\tau; a^*, b^*)$, then simplify, then plug in the true value of $h(\tau)$, then multiply top and bottom by $\left(1 + ce^{-\lambda\tau}\right)$, and then simplify.

Thus a and b are identified.

Since

$$ce^{-\lambda\tau} = \frac{h(\tau) - b}{a - b},$$

c and λ are clearly identified once a and b have been identified.

D.2 Proof of Theorem 1

We show this in three pieces. We first show identification of λ_A^n and λ_B^n, then $P^*, \delta, \lambda_B^e, \lambda_A^e$, and $P(BA)/P(AB)$, and finally $P(AB), P(BA), P(0), P(A0)$, and $P(B0)$.

Identification of λ_A^n and λ_B^n

Condition on workers with the following job history: they are initially non-employed, start at either type firm, become non-employed, start at a B type firm, become non-employed, and then start at an A type firm. We know these are either AB types or BA types. The probability that the first firm is a B type firm is

$$P_B \equiv \frac{\lambda_B^n}{\lambda_A^n + \lambda_B^n},$$

[1] To see this note that the derivative of the log will depend on τ otherwise while at $(a^*, b^*) = (a, b)$ it would be constant.

and this can be directly identified from the data.

We define P_A in an analogous manner. Let T_1 be the duration of the first non-employment spell, T_2 be the duration of the first employment spell, T_3 the duration of the second non-employment spell, T_4 the duration of the second employment spell, and T_5 the duration of the third non-employment spell. For any values $\left(t_1^a, t_1^b, t_2, t_3, t_4, t_5\right)$ we can identify

$$\frac{Pr\left(T_1 > t_1^a, T_2 \leq t_{2,} T_3 \leq t_3, T_4 \leq t_4, T_5 \leq t_5\right)}{Pr\left(T_1 > t_1^b, T_2 \leq t_{2,} T_3 \leq t_3, T_4 \leq t_4, T_5 \leq t_5\right)} = \frac{e^{-\left(\lambda_A^n + \lambda_B^n\right)t_1^a}}{e^{-\left(\lambda_A^n + \lambda_B^n\right)t_1^b}}.$$

Clearly, $\lambda_A^n + \lambda_B^n$ is identified as long as we pick values such that $t_1^a \neq t_1^b$. From this sum and the definition of P_B we can identify λ_A^n and λ_B^n. Note, that we need to worry about $t_2, ..., t_5$ because we condition on people who experience both A and B type firms before the period ends.

Identification of $P^*, \delta, \lambda_B^e, \lambda_A^e$, **and** $P(BA)/P(AB)$

First, we establish identification of P^*. This is simplified by continuing to condition on people for whom we know either $C_i = AB$ or $C_i = BA$. We do this by conditioning on individuals who start at a type B job, leave to non-employment and then start a type A job at t_1. We derive the hazards of those moving from A to B and from A to non-employment at time $t_1 + \tau$. Since they have accepted offers from both jobs we know that $C_i \in \{AB, BA\}$. The events that put one into this conditioning set are independent of type, so the relative proportion of $C_i = AB$ versus to $C_i = BA$ will be the same as it is in the population.

The hazard rate out of job A and into job B is

$$\delta P^* P_B.$$

for the AB types and

$$\delta P^* P_B + \lambda_B^e.$$

for the BA types. We can write the unconditional hazard rate out of job A and into job B as a weighted average of the two where the weights are the survivor functions in job A

$$\frac{P\left(AB\right)e^{-\delta[1-P^*P_A]\tau}\delta P^* P_B + P\left(BA\right)e^{-\lambda_B^e\tau}e^{-\delta[1-P^*P_A]\tau}\left[\delta P^* P_B + \lambda_B^e\right]}{P\left(AB\right)e^{-\delta[1-P^*P_A]\tau} + P\left(BA\right)e^{-\lambda_B\tau}e^{-\delta[1-P^*P_A]\tau}}.$$

Note first, that if $P\left(AB\right) = 0$ or $P\left(BA\right) = 0$ the hazard will be constant, otherwise it will not be (since $\lambda_B^e > 0$). Thus if the hazard is not constant we know $P(A\dot{B}) > 0$ and $P(BA) > 0$. If it is constant we know either $P\left(AB\right) = 0$ or $P\left(BA\right) = 0$, but not distinguish between these two cases from the hazard rate above.

In the case in which both $P(AB)$ and $P(BA)$ are greater than zero, this is a special case of the hazard in Lemma D.1 above with

$$a = \delta P^* P_B$$
$$b = \delta P^* P_B + \lambda_B^e$$
$$c = \frac{P(BA)}{P(AB)}$$
$$\lambda = \lambda_B^e.$$

This allows us to identify λ_B^e, $P(BA)/P(AB)$, and δP^* (since P_B is identified)

We can also identify the conditional hazard of movements to non-employment at time $t_1 + \tau$. This is

$$\frac{P(AB) e^{-\delta[1-P^* P_A]\tau} \delta [1 - P^*] + P(BA) e^{-\lambda_B^e \tau} e^{-\delta[1-P^* P_A]\tau} \delta [1 - P^*]}{P(AB) \int e^{-\delta[1-P^* P_A]\tau} + P(BA) e^{-\lambda_B^e \tau} e^{-\delta[1-P^* P_A]\tau}} = \delta [1 - P^*].$$

Thus, $\delta = \delta [1 - P^*] + \delta P^*$ is identified as is P^*.

Since the model is symmetric we can use the analogous argument to identify λ_A^e.

Next, consider the case in which either $P(AB) = 0$ or $P(AB) = 0$. If $P(AB) = 0$ the hazard rate from A to B will be $\delta P^* P_B$ and the hazard rate from B to A will be $\delta P^* P_A + \lambda_B^e$. If $P(BA) = 0$ the hazard rate from A to B will be $\delta P^* P_B + \lambda_A^e$ and the rate from B to A will be $\delta P^* P_A$. Note that the ratio of the second to the first will be

$$\frac{\delta P^* P_A + \lambda_B^e}{\delta P^* P_B} > \frac{P_A}{P_B},$$

in the first case and

$$\frac{\delta P^* P_A}{\delta P^* P_B + \lambda_A^e} < \frac{P_A}{P_B},$$

in the second case, so these are separately identified. In the first case we can identify P^* and λ_B^e, while in the second we can identify P^* and λ_A^e.

Identification of $P(AB), P(BA), P(0)$, $P(A0)$, and $P(B0)$

First, consider the hazard rate into non-employment for people who start at an A job in their first job. This is similar to the case above except we no longer condition on having held a B job. This now includes three groups BA, AB, and $A0$.

To identify the $P(A0)$ group we use the same argument we used for identification of P^* except that we no longer condition on having a B spell prior to the A spell. We condition on all individuals who's first spell is a type A spell which starts at t_1 and condition on how it ends. Now three preference groups can experience the A spell: AB,BA, and $A0$. The hazard of moving from job A to job B at time τ is

$$\begin{aligned}
&\frac{P(AB)\, e^{-\delta[1-P^*P_A]\tau}\delta P^* P_B + P(BA)\, e^{-\lambda_B^e\tau}e^{-\delta[1-P^*P_A]\tau}\left[\delta P^* P_B + \lambda_B^e\right]}{P(AB)\, e^{-\delta[1-P^*P_A]\tau} + P(BA)\, e^{-\lambda_B\tau}e^{-\delta[1-P^*P_A]\tau} + P(A0)\, e^{-\delta[1-P^*P_A]\tau}}\\
=&\frac{e^{-\delta[1-P^*P_A]\tau}\delta P^* P_B + \frac{P(BA)}{P(AB)}e^{-\lambda_B^e\tau}e^{-\delta[1-P^*P_A]\tau}\left[\delta P^* P_B + \lambda_B^e\right]}{e^{-\delta[1-P^*P_A]\tau} + \frac{P(BA)}{P(AB)}e^{-\lambda_B\tau}e^{-\delta[1-P^*P_A]\tau} + \frac{P(A0)}{P(AB)}e^{-\delta[1-P^*P_A]\tau}}.
\end{aligned}$$

We have identified everything in this expression other than $\frac{P(A0)}{P(AB)}$, so it must be identified.

An analogous argument will identify $\frac{P(B0)}{P(AB)}$. To simplify the expression define

$$\begin{aligned}
\rho_{BA} =&\frac{P(BA)}{P(AB)}\\
\rho_{A0} =&\frac{P(A0)}{P(AB)}\\
\rho_{B0} =&\frac{P(B0)}{P(AB)},
\end{aligned}$$

where the values of ρ are all identified. Then we know that

$$1 = P(0) + [1 + \rho_{BA} + \rho_{A0} + \rho_{B0}]\, P(AB),$$

and we can also identify the probability that someone has not found a job yet at time τ which is

$$P(0) + \left[\rho_{A0}e^{-\lambda_A^n\tau} + \rho_{B0}e^{-\lambda_B^n\tau} + (\rho_{BA}+1)\, e^{-\left[\lambda_A^n+\lambda_B^n\right]\tau}\right] P(AB),$$

from which we can easily solve for $P(AB)$ and then the rest of the probability

D.3 Theorem 2

The proof of theorem 2 is very tedious because there are many pieces. To get an intuition for how this proof will work we first consider an even simpler version of our model. We then present the proof.

Intuition using Simplified Model

We simplify to one job, no human capital, no measurement error, and that all workers take the job when it is offered (which means selection is not a problem). There are only two relevant wages, the one received right after non-employment, R_{i0}, and the one received when the worker gets a competitive outside offer. This will be π_i since both firms are willing to pay this wage. Our goal is to estimate the joint distribution of (R_{i0}, π_i), which we do by identifying the joint characteristic functions

$$E\left[exp\left(i\left(s_1 R_{i0} + s_2 \pi_i\right)\right)\right],$$

for any s_1 and s_2. To do this we need two periods of wages measured at time 1 and time 2. The complication is that during each of the two periods we do not know whether the worker has received an outside offer or not (i.e. the second period wage could be R_{i0} or it could be π_i). To solve this problem we condition on people who are working continuously at the same firm in both time periods and we vary the start time of that spell. Specifically, let $1 - d$ be the time at which the job spell starts. There are three possibilities

- The worker receives an outside offer before time 1. This happens with probability $\left(1 - e^{-\lambda d}\right)$
- The worker receives an outside offer between time 1 and 2. This happens with probability $e^{-\lambda d} - e^{-\lambda(d+1)}$
- The worker does not receive an outside offer before time 2. This happens with probability $e^{-\lambda(d+1)}$.

Then we can write the overall characteristic function as

$$\begin{aligned} & E\left[exp\left(i\left(s_1 W_{i1} + s_2 W_{i1}\right)\right) \mid d\right] \\ = & \left(1 - e^{-\lambda d}\right) E\left[exp\left(i\left(s_1 \pi_i + s_2 \pi_i\right)\right)\right] + \left(e^{-\lambda d} - e^{-\lambda(d+1)}\right) E\left[exp\left(i\left(s_1 R_{i0} + s_2 \pi_i\right)\right)\right] \\ & + e^{-\lambda(d+1)} E\left[exp\left(i\left(s_1 R_{i0} + s_2 R_{i0}\right)\right)\right]. \end{aligned}$$

For any s_1 and s_2, we can move d continuously, so we are generally overidentified as we have many equations and only 3 unknowns: $E\left[exp\left(i\left(s_1 \pi_i + s_2 \pi_i\right)\right)\right]$, $E\left[exp\left(i\left(s_1 R_{i0} + s_2 \pi_i\right)\right)\right]$, and $E\left[exp\left(i\left(s_1 R_{i0} + s_2 R_{i0}\right)\right)\right]$. Thus, intuitively this is identified-and it is pretty clear in this case that since $\left(1 - e^{-\lambda d}\right)$ is nonlinear, we should be able to find values of d to identify it as the model seems clearly over-identified.

To show identification formally, we consider some special cases of d. First, consider the first time period only ($s_2 \doteq 0$). Note that

$$\begin{aligned} lim_{d \downarrow 0} E\left[exp\left(i\left(s_1 W_{i1}\right)\right) \mid d\right] = & lim_{d \downarrow 0}\left[\left(1 - e^{-\lambda d}\right) E\left[exp\left(i\left(s_1 \pi_i\right)\right)\right] + e^{-\lambda d} E\left[exp\left(i\left(s_1 R_{i0}\right)\right)\right]\right] \\ = & E\left[exp\left(i\left(s_1 R_{i0}\right)\right)\right] \end{aligned}$$

which is the characteristic function of R_{i0} and it is identified for any value of s_1.

Use this same equation but take any other value of d then

$$E\left[exp\left(i\left(s_1\pi_i\right)\right)\right]=\frac{E\left[exp\left(i\left(s_1W_{i1}\right)\right)\right]-e^{-\lambda d}\psi_R\left(s_1\right)}{\left(1-e^{-\lambda d}\right)},$$

is also identified for any value of s_1.

Finally, note that

$$\begin{aligned}&E\left[exp\left(i\left(s_1R_{i0}+s_2\pi_i\right)\right)\right]\\=&\frac{E\left[exp\left(i\left(s_1W_{i1}+s_2W_{i1}\right)\right)\right]-\left(1-e^{-\lambda d}\right)E\left[exp\left(i\left(s_1\pi_i+s_2\pi_i\right)\right)\right]-e^{-\lambda(d+1)}E\left[exp\left(i\left(s_1R_{i0}+s_2R_{i0}\right)\right)\right]}{\left(e^{-\lambda d}-e^{-\lambda(d+1)}\right)}\\=&\frac{E\left[exp\left(i\left(s_1W_{i1}+s_2W_{i1}\right)\right)\right]-\left(1-e^{-\lambda d}\right)E\left[exp\left(i\left(s_{12}\pi_i\right)\right)\right]-e^{-\lambda(d+1)}E\left[exp\left(i\left(s_{12}R_{i0}\right)\right)\right]}{\left(e^{-\lambda d}-e^{-\lambda(d+1)}\right)}\end{aligned}$$

where $s_{12}\equiv s_1+s_2$. Since everything on the right hand side is identified, the left is as well.

Proof of Theorem 2

To shorten some of the expressions we will use shorthand notation $\varpi_{ij\ell h_0 h}$ which we define as

$$\varpi_{ij\ell h_0 h}\equiv\log\left(R_{ij\ell h_0}\psi_h\right).$$

Identification of Distribution of Measurement Error (ξ_{it})

First, we identify the distribution of measurement error and then the arrival rate of human capital, λ_h. We condition on a group who

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

We assume that the d_j's are sufficiently small, so spells do not overlap.

We can identify the joint distribution of (w_{i1},w_{i2}) conditional on the events above for alternative values of d_1,d_2, and d_3.

Taking limits of the above object as $d_1\downarrow 0, d_2\downarrow 0$, and $d_3\downarrow 0$, we can identify the conditional distribution of

$$\left(\varpi_{iA000}+\xi_{i1},\varpi_{iA000}+\xi_{i2}\right),$$

for our conditioning group. The first component of these wages will correspond to r_{iA00}, because the workers 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 ϖ_{iA000}.

Identification of λ_h

Next, we show identification of λ_h. To economize on notation we will use $E(\cdot \mid d)$ to denote the expectation conditional on the events described above at values of $d = (d_1, d_2, d_3)$. We use the same conditioning group as in the Measurement Error section and continue to send $d_1 \downarrow 0$ and $d_3 \downarrow 0$, but allow d_2 to vary. We can identify the conditional characteristic function

$$\lim_{d_1, d_3 \downarrow 0} \frac{E\left(e^{\imath s w_{i2}} \mid d\right)}{\phi_{\xi}(s)} = \lim_{d_1, d_3 \downarrow 0} \left[e^{-\lambda_h d_2} E\left(e^{\imath s \varpi_{iA000}} \mid d\right) + \left(1 - e^{-\lambda_h d_2}\right) E\left(e^{\imath s \varpi_{iA011}} \mid d\right)\right].$$

By varying d_2 we can identify λ_h.[2] Intuitively, varying d_2 varies the time that the worker has to receive a human capital shock.

Identification of ψ_1, δ, and the Distribution of Wages for the AB types

We now consider identification of ψ_1 and demonstrate identification of the full wage distribution for the AB type. Identification of the latter is complicated, so we will do this in steps by showing identification of expanding subsets of the full distribution.

For the AB types there are the seventeen different labor market statuses possible

[2] To see how, take the ratio of the derivatives of this function in terms of d_2 at two different values of d_2 and it will be a known function of λ_h. First note that the derivative with respect to d_2 is

$$\lim_{d_1, d_3 \downarrow 0} \left[-\lambda_h e^{-\lambda_h d_2} E\left(e^{\imath s r_{iA00}} \mid d\right) + \lambda_h e^{-\lambda_h d_2} E\left(e^{\imath s r_{iA01}} \mid d\right)\right] = \lambda_h e^{-\lambda_h d_2} \left[E\left(e^{\imath s r_{iA00}} \mid A\right) - E\left(e^{\imath s r_{iA01}} \mid A\right)\right]$$

where the notation $E(\cdot \mid A)$ means the expected value conditional on taking an A job first. Now take the ratio of this at two different values of d_2 say d_2^a and d_2^b then

$$\begin{aligned}\Delta(d_2^a, d_2^b) \equiv& \frac{\lambda_h e^{-\lambda_h d_2^a} \left[E\left(e^{\imath s r_{iA00}} \mid A\right) - E\left(e^{\imath s r_{iA01}} \mid A\right)\right]}{\lambda_h e^{-\lambda_h d_2^b} \left[E\left(e^{\imath s r_{iA00}} \mid A\right) - E\left(e^{\imath s r_{iA01}} \mid A\right)\right]} \\ =& e^{\lambda_h \left(d_2^b - d_2^a\right)}.\end{aligned}$$

$\Delta(d_2^a, d_2^b)$ is directly identified from the data and

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

Table D1
Labor Market Statuses for AB workers

$j(i,t)$	$h(i,t)$	$\ell(i,t)$	$h_0(i,t)$	Wage	Log(Wage)
A	0	0	0	R_{iA00}	ω_{iA000}
A	0	B	0	R_{iAB0}	ω_{iAB00}
A	0	A	0	π_{iA}	ω_{iAA00}
A	1	0	0	$R_{iA00}\psi_1$	ω_{iA001}
A	1	B	0	$R_{iAB0}\psi_1$	ω_{iAB01}
A	1	A	0	$\pi_{iA}\psi_1$	ω_{iAA01}
A	1	0	1	$R_{iA01}\psi_1$	ω_{iA011}
A	1	B	1	$R_{iAB1}\psi_1$	ω_{iAB11}
A	1	A	1	$\pi_{iA}\psi_1$	ω_{iAA11}
B	0	0	0	R_{iB00}	ω_{iB000}
B	0	B	0	π_{iB}	ω_{iBB00}
B	1	0	0	$R_{iB00}\psi_1$	ω_{iB001}
B	1	B	0	$\pi_{iB}\psi_1$	ω_{iBB01}
B	1	0	1	$R_{iB01}\psi_1$	ω_{iB011}
B	1	B	1	$\pi_{iB}\psi_1$	ω_{iBB11}
0	0	NA	NA	NA	NA
0	1	NA	NA	NA	NA

where as a reminder $j(i,t)$ is the current job, $h(i,t)$ is the current level of human capital, and $\ell(i,t)$ and $h_0(i,t)$ are respectively the outside option and level of human capital when the current hmuna capital rental rate was negotiated.

From Table D1 one can see that for an AB worker's wages depend on the joint distribution of eight objects (in addition to ψ_1)

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

The model is overidentified so there are multiple ways to show identification. We focus on a particular set of transitions and show identification by taking limits. We emphasize that this is sufficient to show identification, we do not think it is necessary. We assume that workers start their labor market career in non-employed and receive their first job at $1 - d_1$. The following table shows the transition path.

Transition	Time
Start at A	$1 - d_1$
Move to non-employment	$1 + d_2$
Start at B	$2 - d_3$
Move to non-employment	$2 + d_4$
Start at B	$3 - d_5 - d_6$
Move to A	$3 - d_6$
Move to non-employment	$3 + d_7$
Start at A	$4 - d_8$
Move to non-employment	$4 + d_9$
Start at B	$6 - d_{10}$
Move to non-employment	$6 + d_{11}$
Start at B	$8 - d_{12} - d_{13}$
Move to A	$8 - d_{13}$
Move to B	$8 + d_{14}$

with $d_j \geq 0$ for $j = 1, ..., 14$. We assume that the d_j's are sufficiently small such that the above spells do not overlap. The goal here will be to look at the joint distribution of wages conditional on the d_j's. Analogous to above, we use the notation $E[\cdot \mid d]$ to mean the conditional expectation conditioning on events occurring at times denoted by $d_1 - d_{14}$.

Identification of Distribution of $(w_{i1}, ..., w_{i8})$ conditional on $(d, C_i = AB)$.

In going forward we condition on functions of wages from the first eight periods $f(w_{i1}, ..., w_{i8})$. Since we observe these workers at both firm types, we know they are either AB types or BA types. For any function $f(w_{i1}, ..., w_{i8})$ notice that

$$
\begin{aligned}
E\left(f(w_{i1}, ..., w_{i8}) \mid d\right) =& P\left[AB \mid d\right] E\left[f(w_{i1}, ..., w_{i8}) \mid d, AB\right] \\
&+ P\left[BA \mid d\right] E\left[f(w_{i1}, ..., w_{i8}) \mid d, BA\right].
\end{aligned}
$$

The last value d_{14} will play a crucial roll in distinguishing between these expressions. As it is not realized until after period 8 it does not affect either $E\left[f(w_{i1}, ..., w_{i8}) \mid d, AB\right]$ or $E\left[f(w_{i1}, ..., w_{i8}) \mid d, BA\right]$. However, it does influence $P\left[AB \mid d\right]$, because a BA type can move from A to B directly either because they got an outside offer from a B firm or because the they were laid off and got an immediate offer. For an AB type it can only be due to the latter event. It is straight forward to show that given the result of Theorem 1, $P\left[AB \mid d\right]$ is a known function of d.[3]

[3] It is

$$
P\left[AB \mid d\right] = \frac{a}{a+b},
$$

where

$$
\begin{aligned}
a &= P\left(AB\right) E\left(e^{-(\lambda_A^e + \delta_i)(d_5 + d_{12}) - \delta_i(d_{13} + d_{14})} \left[\delta_i P^* P_A + \lambda_A^e\right]^2 \delta_i P^* P_B\right) \\
b &= P(BA) E\left(e^{-\delta_i(d_5 + d_{12}) - (\lambda_B^e + \delta_i)(d_{13} + d_{14})} \left[\delta_i P^* P_A\right]^2 \left[\delta_i P^* P_B + \lambda_B^e\right]\right).
\end{aligned}
$$

Notice then for any distribution of wages, $f(\cdot)$, and any values of $d_1, ..., d_{13}$, by moving d_{14} we can separately identify $E[f(w_{i1}, ..., w_{i8}) \mid d, AB]$ from $E[f(w_{i1}, ..., w_{i8}) \mid d, BA]$. We refrain from making this argument repeatedly but just condition on types implicitly assuming that $E[f(w_{i1}, ..., w_{i8}) \mid d, AB]$ is identified.

While in principle we could show full identification of the eight dimensional distribution all at once, it is very complicated so instead we show it in pieces. We start with 3 parts.

Identification of joint distribution of $(R_{iA00}, R_{iB00}, R_{iAB0})$ for the AB types

We start by sending $d_1 ... d_6 \downarrow 0$ and look at the joint distribution of (w_{i1}, w_{i2}, w_{i3}). A complication is that at time $3 - d_6$ individuals who moved directly from B to A could have either gotten an outside offer from an A firm or been laid off and found a new job immediately. Define $\rho_3(d)$ to be the probability that it is a voluntary transition. This a complicated but known expression since it involves only transition parameters which we showed are identified in Theorem 1.

Then for any values of $s_1 - s_3$ we can identify

$$\begin{aligned}
&\lim_{d_1 ..., d_6 \downarrow 0} \frac{E[\exp(i(s_1 w_{i1} + s_2 w_{i2} + s_3 w_{i3})) \mid d, AB]}{\phi_\xi(s_1)\phi_\xi(s_2)\phi_\xi(s_3)} \\
= &\left[\lim_{d_1 ..., d_6 \downarrow 0} \rho_3(d)\right] E[\exp(i(s_1 \varpi_{iA000} + s_2 \varpi_{iB000} + s_3 \varpi_{iAB00})) \mid AB] \\
&+ \left[\lim_{d_1 ..., d_6 \downarrow 0} (1 - \rho_3(d))\right] E[\exp(i((s_1 + s_3)\varpi_{iA000} + s_2 \varpi_{iB000})) \mid AB]. \qquad \text{(E.2.1)}
\end{aligned}$$

We will use the same basic argument for identification of the model throughout this section. We will be explicit about it here, but not as explicit in what follows (which will involve many more terms).

1. $\lim_{d_1 ..., d_6 \downarrow 0} \rho_3(d)$ is identified as it is a known function of parameters that we have shown are identified.

2. By setting $s_3 = 0$ we can identify $E[\exp(i(s_1 \varpi_{iA000} + s_2 \varpi_{iB000})) \mid AB]$ from the expression above.

3. Once this is identified, $E[\exp(i((s_1 + s_3)\varpi_{iA000} + s_2 \varpi_{iB000})) \mid AB]$ is identified as we vary s_3.

4. Everything in the expression (E.2.1) above is identified except

   $E[\exp(i(s_1 \varpi_{iA000} + s_2 \varpi_{iB000} + s_3 \varpi_{iAB000})) \mid AB]$ so we can solve for this expression as well.

5. $E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}\right)\right)\mid AB\right]$ is the characteristic function of $\left(\varpi_{iA000},\varpi_{iB000},\varpi_{iAB00}\right)$, so since this is identified, the joint distribution of $\left(R_{iA00},R_{iB00},R_{iAB0}\right)$ conditional on $C_i=AB$ is identified.

Identification of $\left(R_{iA00},R_{iB00},R_{iAB0},R_{A01},R_{iB01},R_{iAB1}\right)$ for the AB types

Now we extend the argument to include the joint distribution of

$$\left(R_{iA00},R_{iB00},R_{iAB0},R_{A01},R_{iB01},R_{iAB1}\right),$$

for the AB types by adding wages for periods 4, 6, and 8. We will now vary d_7 which will allow for the possibility that human capital evolves between time 3 and $3+d_7$ but send other values of d towards 0. There are 8 possible indistinguishable events that can occur in the data; (the job-to-job transition to job A at time $3-d_6$ is voluntary/involuntary) $\times$(human capital evolves or does not evolve between period 3 and $3+d_7$)$\times$ (the job-to-job transition to job A at time $8-d_{13}$ is voluntary/involuntary). Let ρ_3 and ρ_8 be the limit as $d_1...,d_6,d_8,...,d_{13}\downarrow 0$ of the conditional probability that the job-to-job transitions are voluntary at time $3-d_6$ and $8-d_{13}$, respectively. These are identified as they depend on transition parameters that we have shown are identified.

For any value of s_1-s_6 we can identify

$$\begin{aligned}
&\lim_{d_1...,d_6,d_8,...,d_{13}\downarrow 0}\frac{E\left[\exp\left(i\left(s_1w_{i1}+s_2w_{i2}+s_3w_{i3}+s_4w_{i4}+s_5w_{i6}+s_6w_{i8}\right)\right)\mid d,AB\right]}{\phi_\xi\left(s_1\right)\phi_\xi\left(s_2\right)\phi_\xi\left(s_3\right)\phi_\xi\left(s_4\right)\phi_\xi\left(s_5\right)\phi_\xi\left(s_6\right)}\\
=&e^{-\lambda_h d_7}\left[\rho_3\rho_8\right]E\left[\exp\left(i\left(\left(s_1+s_4\right)\varpi_{iA000}+\left(s_2+s_5\right)\varpi_{iB000}+\left(s_3+s_6\right)\varpi_{iAB00}\right)\right)\mid AB\right]\\
&+e^{-\lambda_h d_7}\left[\rho_3\left(1-\rho_8\right)\right]E\left[\exp\left(i\left(\left(s_1+s_4+s_6\right)\varpi_{iA000}+\left(s_2+s_5\right)\varpi_{iB000}+s_3\varpi_{iAB00}\right)\right)\mid AB\right]\\
&+e^{-\lambda_h d_7}\left[\left(1-\rho_3\right)\rho_8\right]E\left[\exp\left(i\left(\left(s_1+s_3+s_4\right)\varpi_{iA000}+\left(s_2+s_5\right)\varpi_{iB000}+s_6\varpi_{iAB00}\right)\right)\mid AB\right]\\
&+e^{-\lambda_h d_7}\left[\left(1-\rho_3\right)\left(1-\rho_8\right)\right]E\left[\exp\left(i\left(\left(s_1+s_3+s_4+s_6\right)\varpi_{iA000}+\left(s_2+s_5\right)\varpi_{iB000}\right)\right)\mid AB\right]\\
&+\left(1-e^{-\lambda_h d_7}\right)\left[\rho_3\rho_8\right]E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+s_4\varpi_{iA011}+s_5\varpi_{iB011}+s_6\varpi_{iAB11}\right)\right)\mid AB\right]\\
&+\left(1-e^{-\lambda_h d_7}\right)\left[\rho_3\left(1-\rho_8\right)\right]E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+\left(s_4+s_6\right)\varpi_{iA011}+s_5\varpi_{iB011}\right)\right)\mid AB\right]\\
&+\left(1-e^{-\lambda_h d_7}\right)\left[\left(1-\rho_3\right)\rho_8\right]E\left[\exp\left(i\left(\left(s_1+s_3\right)\varpi_{iA000}+s_2\varpi_{iB000}+s_4\varpi_{iA011}+s_5\varpi_{iB011}+s_6\varpi_{iAB11}\right)\right)\mid AB\right]\\
&+\left(1-e^{-\lambda_h d_7}\right)\left[\left(1-\rho_3\right)\left(1-\rho_8\right)\right]E\left[\exp\left(i\left(\left(s_1+s_3\right)\varpi_{iA000}+s_2\varpi_{iB000}+\left(s_4+s_6\right)\varpi_{iA011}+s_5\varpi_{iB011}\right)\right)\mid AB\right].
\end{aligned}$$

We showed above that the first four expressions are identified. Thus, we have four new expressions to identify:

$$(a) E\exp\left[i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+s_4\varpi_{iA011}+s_5\varpi_{iB011}+s_6\varpi_{iAB11}\right)\mid AB\right]$$
$$(b) E\exp\left[i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+\left(s_4+s_6\right)\varpi_{iA011}+s_5\varpi_{iB011}\right)\mid AB\right]$$
$$(c) E\exp\left[i\left(\left(s_1+s_3\right)\varpi_{iA000}+s_2\varpi_{iB000}+s_4\varpi_{iA011}+s_5\varpi_{iB011}+s_6\varpi_{iAB11}\right)\mid AB\right]$$
$$(d) E\exp\left[i\left(\left(s_1+s_3\right)\varpi_{iA000}+s_2\varpi_{iB000}+\left(s_4+s_6\right)\varpi_{iA011}+s_5\varpi_{iB011}\right)\mid AB\right]$$

We use the same approach as above. If we evaluate at $s_3 = s_6 = 0$ these expressions are the same and thus $E\left[i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_4\varpi_{iA011}+s_5\varpi_{iB011}\mid AB\right)\right]$ is identified. This identifies (d) for any values of $s_1 - s_6$.

Again, using the same type of argument, given (d), keeping $s_3 = 0$ but varying the other values of s_j identifies (c), and setting $s_6 = 0$ but varying the others gives (b). Then everything in the large equation above is identified other than (a), so it is identified by varying all values of s_j.

Identification of ψ_1

Next, we consider identification of ψ_1 which we can do from $E\left[\log\left(W_{i7}\right)\mid d, AB\right]$ alone. In order to do this we condition on $1 < d_{11} < 2$, so that we observe w_{i7} and we will vary d_2 but send the rest of the $d_j \downarrow 0$ (except d_{14} as usual). For this case there are three possibilities: human capital has not evolved before period 7, human capital evolves between time 1 and time $1 + d_2$, and human capital evolves between periods 6 and 7.

$$\begin{aligned}
&\lim_{d_1,d_3,\ldots,d_{13}\downarrow 0} E\left[\log\left(w_{i7}\right)\mid d, AB\right] = \\
&e^{-\lambda_h[1+d_2]}\left(e^{-\lambda_B^e}E\left[\varpi_{iB000}\mid AB\right]+\left(1-e^{-\lambda_B^e}\right)E\left[\varpi_{iBB00}\mid AB\right]\right)\\
&+\left(1-e^{-\lambda_h d_2}\right)\left(e^{-\lambda_B^e}E\left[\varpi_{iB011}\mid AB\right]+\left(1-e^{-\lambda_B^e}\right)\left(E\left[\varpi_{iBB11}\mid AB\right]\right)\right)\\
&+\left(e^{-\lambda_h d_2}-e^{-\lambda_h[1+d_2]}\right)\left(e^{-\lambda_B^e}E\left[\varpi_{iB001}\mid AB\right]+\left(1-e^{-\lambda_B^e}\right)E\left[\varpi_{iBB11}\mid AB\right]\right)+E\left(\xi_{i7}\right)\\
=&e^{-\lambda_h[1+d_2]}\left(e^{-\lambda_B^e}E\left[\varpi_{iB000}\mid AB\right]+\left(1-e^{-\lambda_B^e}\right)E\left[\log\left(\pi_{iB}\right)\mid AB\right]\right)\\
&+\left(1-e^{-\lambda_h d_2}\right)\left(e^{-\lambda_B^e}E\left[\varpi_{iB011}\mid AB\right]+\left(1-e^{-\lambda_B^e}\right)\left(E\left[\log\left(\pi_{iB}\right)+\log\left(\psi_1\right)\mid AB\right]\right)\right)\\
&+\left(e^{-\lambda_h d_2}-e^{-\lambda_h[1+d_2]}\right)\left(e^{-\lambda_B^e}E\left[\varpi_{iB000}+\log\left(\psi_1\right)\mid AB\right]+\left(1-e^{-\lambda_B^e}\right)E\left[\log\left(\pi_{iB}\right)+\log\left(\psi_1\right)\mid AB\right]\right)+E\left(\xi_{i7}\right)\\
=&e^{-\lambda_h d_2}e^{-\lambda_B^e}E\left[\varpi_{iB000}\mid AB\right]+\left(1-e^{-\lambda_h d_2}\right)e^{-\lambda_B^e}E\left[\varpi_{iB011}\mid AB\right]+E\left(\xi_{i7}\right)\\
&+\left(1-e^{-\lambda_A}\right)E\left[\log\left(\pi_{iB}\right)\mid AB\right]+\left[\left(1-e^{-\lambda_A}\right)+\left(e^{-\lambda_h d_2}-e^{-\lambda_h[1+d_2]}\right)\right]\log\left(\psi_1\right).
\end{aligned}$$

Everything is identified in this expression except $E\left[\log\left(\pi_{iB}\right)\mid AB\right]$ and $\log\left(\psi_1\right)$, so by varying d_2 they can be separately identified.

Identification of $\left(R_{iA00}, R_{iAB0}, R_{iB00}, R_{iA01}, R_{iAB1}, R_{iB01}, \pi_{iA}, \pi_{iB}\right)$ conditional on AB

Now we assume that $1 < d_{11} < 2$ and $1 < d_9 < 2$, so that we observe wages at all times 1,...,8. Using an analogous argument to the discussion of identification conditional on AB

above, by varying d_7, we can identify the expected value of $f(w_{i1}, ..., w_{i8})$ conditional on d and human capital arriving between time 3 and $3 + d_7$ (write this conditioning as $H_{i4} = 1$). To simplify the notation we define

$$\varpi_{iA} \equiv \log(\pi_{iA})$$
$$\varpi_{iB} \equiv \log(\pi_{iB}).$$

We will send the rest of the d_j's to zero (other than d_7, d_9, d_{11}, and d_{14}). Since we condition on human capital arriving between period 3 and $3 + d_7$, we know that the wage in the first period will be approximately R_{iA00}, the second period R_{iB00}, the fourth R_{iA01}, and the sixth R_{iB01}. As before for the third and the eighth period the wage can take two values depending on whether the job-to-job transition was voluntary or not (R_{iA00} or R_{iAB0} in 3 and R_{iA01}or R_{iAB1} in 8). For period 5 the wage can take 3 values depending on outside offers: either R_{iA01} if no outside offers, R_{iAB1} if an offer from a B type only, or π_{iA} if an offer from an A type. Similarly in period 7 the wage can take 2 values depending on whether there was no outside offer (R_{iB01}) or an outside offer from a B firm (π_{iB}).[4] This gives a total of $2 \times 2 \times 3 \times 2 = 24$ different possibilities.

Analogous to above, we define ρ_3 and ρ_8 be the limit as $d_1, ..., d_6, d_8, d_{10}, d_{12}, d_{13} \downarrow 0$ of the conditional probability that the job-to-job transitions are voluntary at time $3-d_6$ and $8 - d_{13}$, respectively.

Putting this together can identify the complicated expression with the relevant 24 terms.

[4]Since we are considering AB types they could not have gotten an offer from an A firm or they would have left.

$$
\begin{aligned}
&\lim_{d_1,\ldots,d_6,d_8,d_{10},d_{12},d_{13}\downarrow 0} \frac{E\left[\exp\left(i\left(s_1 w_{i1}+s_2 w_{i2}+s_3 w_{i3}+s_4 w_{i4}+s_5 w_{i5}+s_6 w_{i6}+s_7 w_{i7}+s_8 w_{i8}\right)\right) \mid d, AB, H_{i4}=1\right]}{\phi_\xi\left(s_1\right)\phi_\xi\left(s_2\right)\phi_\xi\left(s_3\right)\phi_\xi\left(s_4\right)\phi_\xi\left(s_5\right)\phi_\xi\left(s_6\right)\phi_\xi\left(s_7\right)\phi_\xi\left(s_8\right)}\\
&= e^{-(\lambda_A^e+\lambda_B^e)}e^{-\lambda_B^e}\left[\rho_3\rho_8\right]E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+\left(s_4+s_5\right)\varpi_{iA011}+\left(s_6+s_7\right)\varpi_{iB011}+s_8\varpi_{iAB11}\right)\right) \mid AB\right]\\
&+ e^{-(\lambda_A^e+\lambda_B^e)}\left(1-e^{-\lambda_B^e}\right)\left[\rho_3\rho_8\right]E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+\left(s_4+s_5\right)\varpi_{iA011}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}+s_8\varpi_{iAB11}\right)\right) \mid AB\right]\\
&+ \left(1-e^{-\lambda_A^e}\right)e^{-\lambda_B^e}\left[\rho_3\rho_8\right]E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+s_4\varpi_{iA011}+s_5\varpi_{iAA11}+\left(s_6+s_7\right)\varpi_{iB011}+s_8\varpi_{iAB11}\right)\right) \mid AB\right]\\
&+ \left(1-e^{-\lambda_A^e}\right)\left(1-e^{-\lambda_B^e}\right)\left[\rho_3\rho_8\right]E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+s_4\varpi_{iA011}+s_5\varpi_{iAA11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}+s_8\varpi_{iAB11}\right)\right) \mid AB\right]\\
&+ e^{-\lambda_A^e}\left(1-e^{-\lambda_B^e}\right)e^{-\lambda_B^e}\left[\rho_3\rho_8\right]E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+s_4\varpi_{iA011}+\left(s_6+s_7\right)\varpi_{iB011}+\left(s_5+s_8\right)\varpi_{iAB11}\right)\right) \mid AB\right]\\
&+ e^{-\lambda_A^e}\left(1-e^{-\lambda_B^e}\right)\left(1-e^{-\lambda_B^e}\right)\left[\rho_3\rho_8\right]E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+s_4\varpi_{iA011}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}+\left(s_5+s_8\right)\varpi_{iAB11}\right)\right) \mid AB\right]\\
&+ e^{-(\lambda_A^e+\lambda_B^e)}e^{-\lambda_B^e}\left[\rho_3\left(1-\rho_8\right)\right]E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+\left(s_4+s_5+s_8\right)\varpi_{iA011}+\left(s_6+s_7\right)\varpi_{iB011}\right)\right) \mid AB\right]\\
&+ e^{-(\lambda_A^e+\lambda_B^e)}\left(1-e^{-\lambda_B^e}\right)\left[\rho_3\left(1-\rho_8\right)\right]E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+\left(s_4+s_5+s_8\right)\varpi_{iA011}+s_7\varpi_{iBB11}+s_6\varpi_{iB011}\right)\right) \mid AB\right]\\
&+ \left(1-e^{-\lambda_A^e}\right)e^{-\lambda_B^e}\left[\rho_3\left(1-\rho_8\right)\right]E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+\left(s_4+s_8\right)\varpi_{iA011}+s_5\varpi_{iAA11}+\left(s_6+s_7\right)\varpi_{iB011}\right)\right) \mid AB\right]\\
&+ \left(1-e^{-\lambda_A^e}\right)\left(1-e^{-\lambda_B^e}\right)\left[\rho_3\left(1-\rho_8\right)\right]E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+\left(s_4+s_8\right)\varpi_{iA011}+s_5\varpi_{iAA11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\right)\right) \mid AB\right]\\
&+ e^{-\lambda_A^e}\left(1-e^{-\lambda_B^e}\right)e^{-\lambda_B^e}\left[\rho_3\left(1-\rho_8\right)\right]E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+\left(s_4+s_8\right)\varpi_{iA011}+s_5\varpi_{iAB11}+\left(s_6+s_7\right)\varpi_{iB011}\right)\right) \mid AB\right]\\
&+ e^{-\lambda_A^e}\left(1-e^{-\lambda_B^e}\right)\left(1-e^{-\lambda_B^e}\right)\left[\rho_3\left(1-\rho_8\right)\right]E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+\left(s_4+s_8\right)\varpi_{iA011}+s_5\varpi_{iAB11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\right)\right) \mid AB\right]\\
&+ e^{-(\lambda_A^e+\lambda_B^e)}e^{-\lambda_B^e}\left[\left(1-\rho_3\right)\rho_8\right]E\left[\exp\left(i\left(\left(s_1+s_3\right)\varpi_{iA000}+s_2\varpi_{iB000}+\left(s_4+s_5\right)\varpi_{iA011}+\left(s_6+s_7\right)\varpi_{iB011}+s_8\varpi_{iAB11}\right)\right) \mid AB\right]\\
&+ e^{-(\lambda_A^e+\lambda_B^e)}\left(1-e^{-\lambda_B^e}\right)\left[\left(1-\rho_3\right)\rho_8\right]E\left[\exp\left(i\left(\left(s_1+s_3\right)\varpi_{iA000}+s_2\varpi_{iB000}+\left(s_4+s_5\right)\varpi_{iA011}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}+s_8\varpi_{iAB11}\right)\right) \mid AB\right]\\
&+ \left(1-e^{-\lambda_A^e}\right)e^{-\lambda_B^e}\left[\left(1-\rho_3\right)\rho_8\right]E\left[\exp\left(i\left(\left(s_1+s_3\right)\varpi_{iA000}+s_2\varpi_{iB000}+s_4\varpi_{iA011}+s_5\varpi_{iAA11}+\left(s_6+s_7\right)\varpi_{iB011}+s_8\varpi_{iAB11}\right)\right) \mid AB\right]\\
&+ \left(1-e^{-\lambda_A^e}\right)\left(1-e^{-\lambda_B^e}\right)\left[\left(1-\rho_3\right)\rho_8\right]E\left[\exp\left(i\left(\left(s_1+s_3\right)\varpi_{iA000}+s_2\varpi_{iB000}+s_4\varpi_{iA011}+s_5\varpi_{iAA11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}+s_8\varpi_{iAB11}\right)\right) \mid AB\right]
\end{aligned}
$$

$$
\begin{aligned}
&+ e^{-\lambda_A^e}\left(1-e^{-\lambda_B^e}\right)e^{-\lambda_B^e}\left[\left(1-\rho_3\right)\rho_8\right]E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+\left(s_4+s_5\right)\varpi_{iA011}+\left(s_6+s_7\right)\varpi_{iB011}+s_8\varpi_{iAB11}\right)\right)\mid AB\right]\\
&+ e^{-\lambda_A^e}\left(1-e^{-\lambda_B^e}\right)\left(1-e^{-\lambda_B^e}\right)\left[\left(1-\rho_3\right)\rho_8\right]E\left[\exp\left(i\left(\left(s_1+s_3\right)\varpi_{iA000}+s_2\varpi_{iB000}+s_4\varpi_{iA011}+\left(s_5+s_8\right)\varpi_{iAB11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\right)\right)\mid AB\right]\\
&+ \left(1-e^{-\lambda_A^e}\right)e^{-\lambda_B^e}\left[\left(1-\rho_3\right)\left(1-\rho_8\right)\right]E\left[\exp\left(i\left(\left(s_1+s_3\right)\varpi_{iA000}+s_2\varpi_{iB000}+\left(s_4+s_5+s_8\right)\varpi_{iA011}+\left(s_6+s_7\right)\varpi_{iB011}\right)\right)\mid AB\right]\\
&+ \left(1-e^{-\lambda_B}\right)e^{-\left(\lambda_A+\lambda_B\right)}\left(1-\rho_3\right)\left(1-\rho_8\right)E\left[\exp\left(i\left(\left(s_1+s_3\right)\varpi_{iA000}+s_2\varpi_{iB000}+\left(s_4+s_5+s_8\right)\varpi_{iA011}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\right)\right)\mid AB\right]\\
&+ e^{-\lambda_B}\left(1-e^{-\lambda_A}\right)\left[\left(1-\rho_3\right)\left(1-\rho_8\right)\right]E\left[\exp\left(i\left(\left(s_1+s_3\right)\varpi_{iA000}+s_2\varpi_{iB000}+\left(s_4+s_8\right)\varpi_{iA011}+s_5\varpi_{iAA11}+\left(s_6+s_7\right)\varpi_{iB011}\right)\right)\mid AB\right]\\
&+ \left(1-e^{-\lambda_A^e}\right)\left(1-e^{-\lambda_B^e}\right)\left[\left(1-\rho_3\right)\left(1-\rho_8\right)\right]E\left[\exp\left(i\left(\left(s_1+s_3\right)\varpi_{iA000}+s_2\varpi_{iB000}+\left(s_4+s_8\right)\varpi_{iA011}+s_5\varpi_{iAA11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\right)\right)\mid AB\right]\\
&+ e^{-\lambda_A^e}\left(1-e^{-\lambda_B^e}\right)e^{-\lambda_B^e}\left[\left(1-\rho_3\right)\left(1-\rho_8\right)\right]E\left[\exp\left(i\left(\left(s_1+s_3\right)\varpi_{iA000}+s_2\varpi_{iB000}+\left(s_4+s_8\right)\varpi_{iA011}+s_5\varpi_{iAB11}+\left(s_6+s_7\right)\varpi_{iB011}\right)\right)\mid AB\right]\\
&+ e^{-\lambda_A^e}\left(1-e^{-\lambda_B^e}\right)\left(1-e^{-\lambda_B^e}\right)\left[\left(1-\rho_3\right)\left(1-\rho_8\right)\right]E\left[\exp\left(i\left(\left(s_1+s_3\right)\varpi_{iA000}+s_2\varpi_{iB000}+\left(s_4+s_8\right)\varpi_{iA011}+s_5\varpi_{iAB11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\right)\right)\mid AB\right].
\end{aligned}
$$

We now have sixteen new terms that have not been previously identified.

$(a)E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+(s_4+s_5)\varpi_{iA011}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}+s_8\varpi_{iAB11}\right)\right)\mid AB\right]$

$(b)E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+s_4\varpi_{iA011}+s_5\varpi_{iAA11}+(s_6+s_7)\varpi_{iB011}+s_8\varpi_{iAB11}\right)\right)\mid AB\right]$

$(c)E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+s_4\varpi_{iA011}+s_5\varpi_{iAA11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}+s_8\varpi_{iAB11}\right)\right)\mid AB\right]$

$(d)E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+s_4\varpi_{iA011}+(s_6+s_7)\varpi_{iB011}+(s_5+s_8)\varpi_{iAB11}\right)\right)\mid AB\right]$

$(e)E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+(s_4+s_5+s_8)\varpi_{iA011}+(s_6+s_7)\varpi_{iB011}\right)\right)\mid AB\right]$

$(f)E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+(s_4+s_8)\varpi_{iA011}+s_5\varpi_{iAA11}+(s_6+s_7)\varpi_{iB011}\right)\right)\mid AB\right]$

$(g)E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+(s_4+s_8)\varpi_{iA011}+s_5\varpi_{iAA11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\right)\right)\mid AB\right]$

$(h)E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+(s_4+s_8)\varpi_{iA011}+s_5\varpi_{iAB11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\varpi_{iB}\right)\right)\mid AB\right]$

$(i)E\left[\exp\left(i\left((s_1+s_3)\varpi_{iA000}+s_2\varpi_{iB000}+(s_4+s_5)\varpi_{iA011}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}+s_8\varpi_{iAB11}\right)\right)\mid AB\right]$

$(j)E\left[\exp\left(i\left((s_1+s_3)\varpi_{iA000}+s_2\varpi_{iB000}+s_4\varpi_{iA011}+s_5\varpi_{iAA11}+(s_6+s_7)\varpi_{iB011}+s_8\varpi_{iAB11}\right)\right)\mid AB\right]$

$(k)E\left[\exp\left(i\left((s_1+s_3)\varpi_{iA000}+s_2\varpi_{iB000}+s_4\varpi_{iA011}+s_5\varpi_{iAA11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}+s_8\varpi_{iAB11}\right)\right)\mid AB\right]$

$(l)E\left[\exp\left(i\left((s_1+s_3)\varpi_{iA000}+s_2\varpi_{iB000}+s_4\varpi_{iA011}+(s_5+s_8)\varpi_{iAB11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\right)\right)\mid AB\right]$

$(m)E\left[\exp\left(i\left((s_1+s_3)\varpi_{iA000}+s_2\varpi_{iB000}+(s_4+s_5+s_8)\varpi_{iA011}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\right)\right)\mid AB\right]$

$(n)E\left[\exp\left(i\left((s_1+s_3)\varpi_{iA000}+s_2\varpi_{iB000}+(s_4+s_8)\varpi_{iA011}+s_5\varpi_{iAA11}+(s_6+s_7)\varpi_{iB011}\right)\right)\mid AB\right]$

$(o)E\left[\exp\left(i\left((s_1+s_3)\varpi_{iA000}+s_2\varpi_{iB000}+(s_4+s_8)\varpi_{iA011}+s_5\varpi_{iAA11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\varpi_{iB}\right)\right)\mid AB\right]$

$(p)E\left[\exp\left(i\left((s_1+s_3)\varpi_{iA000}+s_2\varpi_{iB000}+(s_4+s_8)\varpi_{iA011}+s_5\varpi_{iAB11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\right)\right)\mid AB\right]$

We use the same basic approach as above: We set various values of s_j to zero we can identify the components. To see how to identify all of these terms, setting $s_3 = s_8 = s_5 = 0$ all of the terms simplify to either

$$E\exp\left[i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_4\varpi_{iA011}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\right)\mid AB\right]$$

or

$$E\exp\left[i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_4\varpi_{iA011}+(s_6+s_7)\varpi_{iB011}\right)\mid AB\right].$$

However, we have already shown identification of latter of these terms, which means the former is identified. Identification of this gives identification of term (m). Using a similar argument, setting $s_3 = s_8 = s_7 = 0$ we can identify term (n). Given these setting $s_3 = s_5 = 0$ we can show that (p),(i) and (l) are identified. Setting $s_3 = s_8 = 0$ we can identify (o), $s_3 = s_7 = 0$ gives (j), $s_5 = s_8 = 0$ gives (e), and $s_7 = s_8 = 0$ gives (f). Now with these setting $s_3 = 0$ gives (k), $s_8 = 0$ gives (g), $s_5 = 0$ gives (a), (d), and (h), and $s_7 = 0$ gives (b). This leaves only term (c) which is identified by varying all 8 terms given knowledge of all the other terms. This is the characteristic function for the joint distribution. Thus, we have shown that the joint distribution of wages for type AB workers can be non-parametrically identified since the characteristic function uniquely determines the distribution.

Identification of the Distribution of Wages for the Other Types

Using a symmetric argument reversing A and B we can show that the distribution of

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

for the BA types is also identified.

Next consider the $A0$ types. We will use an argument similar to above, though it will be much simpler as there are fewer labor market statuses to worry about.

Table D2
Labor Market Statuses for $A0$ workers

$j(i,t)$	$h(i,t)$	$\ell(i,t)$	$h_0(i,t)$	Wage	$\log(Wage)$
A	0	0	0	R_{iA00}	ω_{iA000}
A	0	A	0	π_{iA}	ω_{iAA00}
A	1	0	0	$R_{iA00}\psi_1$	ω_{iA000}
A	1	A	0	$\pi_{iA}\psi_1$	ω_{iAA01}
A	1	0	1	$R_{iA01}\psi_1$	ω_{iA011}
A	1	A	1	$\pi_{iA}\psi_1$	ω_{iAA11}

From Table D2 one can see that for an $A0$ worker wages depend on the joint distribution of just three objects (in addition to ψ_1)

$$(R_{iA00}, \pi_{iA}, R_{iA01}).$$

Since there are three objects to identify, we only need to use the first three periods. We consider the following the transition path. Individuals begin non-employed at time zero and we will take $d_4 > 1$

Transition	Time
Start at A	$1 - d_1$
Move to non-employment	$1 + d_2$
Start at A	$2 - d_3$
Move to non-employment	$2 + d_4$

We can identify

$$\begin{aligned} E\exp\left[i\left(s_1 w_{i1} + s_2 w_{i2} + s_3 w_{i3}\right) \mid d\right] = & P(AB \mid d) E\exp\left[i\left(s_1 w_{i1} + s_2 w_{i2} + s_3 w_{i3}\right) \mid AB, d\right] \\ & + P(BA \mid d) E\exp\left[i\left(s_1 w_{i1} + s_2 w_{i2} + s_3 w_{i3}\right) \mid BA, d\right] \\ & + P(A0 \mid d) E\exp\left[i\left(s_1 w_{i1} + s_2 w_{i2} + s_3 w_{i3}\right) \mid A0, d\right]. \end{aligned}$$

Since everything else in this expression is identified, we can identify

$$E\exp\left[i\left(s_1 w_{i1} + s_2 w_{i2} + s_3 w_{i3}\right) \mid A0, d\right].$$

Furthermore, analogous to the argument above using d_7, we now vary d_2 to identify the expected value of $f(w_{i1}, ..., w_{i3})$ conditional on d and human capital arriving between time 1 and $1 + d_2$ (write this conditioning as $H_{i2} = 1$). Then we can identify

$$E\exp\left[i\left(s_1 w_{i1} + s_2 w_{i2} + s_3 w_{i3}\right) \mid A0, d, H_{i2} = 1\right].$$

In this case there is only one thing to worry about-whether the worker received and offer from another A firm between periods 2 and 3. Thus, taking $d_1 \downarrow 0$ and $d_3 \downarrow 0$ we can identify

$$\begin{aligned}
&\lim_{d_1,d_3 \downarrow 0} \frac{E \exp\left[i\left(s_1 w_{i1} + s_2 w_{i2} + s_3 w_{i3}\right) \mid A0, d, \delta, H_{i2} = 1\right]}{\phi_\xi\left(s_1\right) \phi_\xi\left(s_2\right) \phi_\xi\left(s_3\right)} \\
=&e^{-\lambda_A} E \exp\left[i\left(s_1 \varpi_{iA000} + \left[s_2 + s_3\right] \varpi_{iA011}\right) \mid A0\right] \\
&+ \left(1 - e^{-\lambda_A}\right) E \exp\left[i\left(s_1 \varpi_{iA000} + s_2 \varpi_{iA011} + s_3 \varpi_{iAA11}\right) \mid A0\right].
\end{aligned}$$

Set $s_3 = 0$ and we can identify $E \exp\left[i\left(s_1 \varpi_{iA000} + s_2 \varpi_{iA011}\right) \mid A0, \delta\right]$. Knowledge of this gives knowledge of $E \exp\left[i\left(s_1 \varpi_{iA000} + \left[s_2 + s_3\right] \varpi_{iA011}\right) \mid A0, \delta\right]$ and then allowing s_3 to vary means we can identify $E \exp\left[i\left(\varpi_{iA000} + s_2 \varpi_{iA011} + s_3 \varpi_{iAA11}\right) \mid A0, \delta\right]$ and thus the joint distribution of $\left(R_{iA00}, \pi_{iA}, R_{iA01}\right)$ conditional on δ for $C_i = A0$.

An analogous argument gives identification of the joint distribution of $\left(R_{iB00}, \pi_{iB}, R_{iB01}\right)$ conditional on δ for $C_i = B0$.

Thus, we have shown that wages, turnover parameters, and type proportions are identified.

D.4 Proof of Theorem 3

The proof here is general enough to cover both the homogeneous δ/finite time case and the heterogeneous δ_i/infinite time case.

In Assumption 4, we have assumed that $Pr(C_i = AB) + Pr(C_i = BA) > 0$, so there are at least one of these two groups. The AB and BA types are symmetric with each other as are the $A0$ and $B0$ types-so we only show the results for the $A0$ and AB types with the $B0$ and BA being analogous. This proof is done in four steps. The first two steps focus on the first part of Theorem 3, where we consider the case in which wages are not bargained (or where workers are indifferent, so bargaining does not matter). First, we do this for $C_i = A0$ and then for $C_i = AB$. The final two steps show that β is not identified. Likewise, we first show this for $C_i = A0$ and then for $C_i = AB$.

We will continue to use the notation from Theorem 2 that

$$\varpi_{ij\ell h_0 h} \equiv \log\left(R_{ij\ell h_0} \psi_h\right)$$

First part of proof for $C_i = A0$

For this group there are four potential wages $\left(R_{iA00}, R_{iAA0}, R_{iA01}, R_{iAA1}\right)$. However, R_{iAA0} and R_{iAA1} are trivially equal to π_{iA}, so only the two relevant endogenous wages are R_{iA00} and

R_{iA01}. They are determined by the following conditions

$$\begin{aligned}
(\rho + \delta_i + \lambda_A^e + \lambda_h)\left[\beta V_{iA0}(\pi_{iA}) + (1-\beta)\, V_{i00}\right] \\
= & u_{iA}\left(R_{iA00}\right) + \lambda_A^e V_{iA0}(\pi_{iA}) \\
& + \delta V_{i00}^* + \lambda_h V_{iA1}\left(R_{iA00}\right)
\end{aligned} \tag{E.3.1}$$

$$\begin{aligned}
(\rho + \delta_i + \lambda_A^e) V_{iA1}\left(R_{iA00}\right) \\
= & u_{iA}\left(R_{iA00}\psi_1\right) + \lambda_A^e V_{iA1}(\pi_{iA}\psi_1) + \delta V_{i01}^*
\end{aligned} \tag{E.3.2}$$

$$\begin{aligned}
(\rho + \delta + \lambda_A^e)\left[\beta V_{iA1}(\pi_{iA}) + (1-\beta)\, V_{i01}\right] \\
= & u_{iA}\left(R_{iA01}\psi_1\right) + \lambda_A^e V_{iA1}(\pi_{iA}) + \delta V_{i01}^*,
\end{aligned} \tag{E.3.3}$$

and we also know that $V_{iA0}(\pi_{iA})$ and $V_{iA1}(\pi_{iA})$ are determined by

$$(\rho + \delta_i + \lambda_h)\, V_{iA0}(\pi_{iA}) = u_{iA}\left(\pi_{iA}\right) + \delta_i V_{i00}^* + \lambda_h V_{iA1}(\pi_{iA}) \tag{E.3.4}$$

$$(\rho + \delta)\, V_{iA1}(\pi_{iA}) = u_{iA}\left(\pi_{iA}\psi_1\right) + \delta_i V_{i01}^*. \tag{E.3.5}$$

Using equations (E.3.1), (E.3.2), (E.3.4), and (E.3.5) through algebra one can show

$$\begin{aligned}
(\rho + \delta_i + \lambda_A^e + \lambda_h)(1-\beta)\left[V_{iA0}(\pi_{iA}) - V_{i00}\right] = & \left[u_{iA}\left(\pi_{iA}\right) - u_{iA}\left(R_{iA00}\right)\right] \\
& + \left[u_{iA}\left(\pi_{iA}\psi_1\right) - u_{iA}\left(R_{iA00}\psi_1\right)\right]\frac{\lambda_h}{\rho + \delta_i + \lambda_A^e}.
\end{aligned} \tag{E.3.6}$$

From which we can see that: (a) if $\beta = 1$ then the left hand side of (E.3.6) is zero so $R_{iA00} = \pi_{iA}$, because utility is strictly increasing in wages. (b) if $V_{i00} = V_{iA0}(\pi_{iA})$ then the left hand side of (E.3.6) is also zero so $R_{iA00} = \pi_{iA}$. (c) if $\beta < 1$ and $V_{iA0}(\pi_{iA}) > V_{i00}$ then the left hand side of (E.3.6) is positive, so the right hand side must be as well. This implies that $R_{iA00} < \pi_{iA}$.

This means that if $\beta = 1$ then $R_{iA00} = \pi_{iA}$ with probability one. Thus, $R_{iA00} < \pi_{iA}$ implies that $\beta < 1$. Secondly, if $R_{iA00} = \pi_{iA}$ with probability one then either $\beta = 1$ or $V_{i00} = V_{iA0}(\pi_{iA})$ with probability one.

Similarly, from equations (E.3.3) and (E.3.5) we can show

$$(\rho + \delta_i + \lambda_A^e + \lambda_h)(1-\beta)\left[V_{iA1}(\pi_{iA}) - V_{i01}\right] = u_{iA}\left(\pi_{iA00}\psi_1\right) - u_{iA}\left(R_{iA01}\psi_1\right). \tag{E.3.7}$$

This is equivalent so if $\beta = 1$ then $R_{iA01} = \pi_{iA}$ with probability one. Thus $R_{iA01} < \pi_{iA}$ implies that $\beta < 1$. If $R_{iA00} = \pi_{iA}$ then either $\beta = 1$ or $V_{i01} = V_{iA1}(\pi_{iA})$ with probability one.

This completes the first part of the proof for $C_i = A0$.

First part of proof for $C_i = AB$

This has very much the same structure as the first proof, but is more complicated in that now there are 6 different endogenous potential wages $R_{iB00}, R_{iB01}, R_{iA00}, R_{iA01}, R_{iAB0}$ and R_{iAB1}.[5]

Four of these wages are straight forward to deal with $(R_{iB00}, R_{iB01}, R_{iA01}, R_{iAB1})$. Going through similar algebra as for the $C_i = A0$ case, it is straight forward to show

$$
\begin{aligned}
(\rho+\delta+\lambda_B^e+\lambda_A^e)(1-\beta)\left[V_{iB1}(\pi_{iB})\right] &= u_{iB}(\pi_{iB}\psi_1) - u_{iB}(R_{iB01}\psi_1) \\
(\rho+\delta+\lambda_B^e+\lambda_A^e)(1-\beta)\left[V_{iA1}(\pi_{iA})\right] &= u_{iA}(\pi_{iA}\psi_1) - u_{iA}(R_{iA01}\psi_1) + \lambda_B^e(1-\beta)\left[V_{iA1}(\pi_{iA}) - V_{iB1}(\pi_{iB})\right] \\
(\rho+\delta+\lambda_A^e)(1-\beta)\left[V_{iA1}(\pi_{iA})\right] &= u_{iA}(\pi_{iA}\psi_1) - u_{iA}(R_{iAB1}\psi_1) \\
(\rho+\delta+\lambda_A+\lambda_h)(1-\beta)\left[V_{iB0}(\pi_{iA}) - V_{i00}\right] &= u_{iB}(\pi_{iB}) - u_{iB}(R_{iB00}) + \lambda_h \frac{u_{iB}(\pi_{iB}\psi_1) - u_{iB}(R_{iB00}\psi_1)}{(\rho+\delta+\lambda_B^e+\lambda_A^e)}.
\end{aligned}
$$

So using the same argument as for the $C_i = A0$ case, since utility is strictly increasing in wages, when $\beta = 1$ then the left hand side of all of these equations is zero, so $R_{iB01} = \pi_{iB}$, $R_{iA01} = \pi_{iA}$, $R_{iAB1} = \pi_{iA}$ and $R_{iB00} = \pi_{iB}$. Thus if any of these equalities does not hold with positive probability, then $\beta < 1$. Finally, if these equalities hold then either $\beta = 1$ or workers are indifferent between all relevant outcomes.

The R_{iAB0} and R_{iA00} cases are more complicated because the shape of the utility function dictates whether the worker would use an offer from a B firm to renegotiate the wage after their human capital augments. We will go over them in more detail.

- Rental rate is R_{iA00} and worker has augmented human capital and an offer from B: In general we would expect that since B is preferred to non-employment one would prefer the B job, but it depends on the utility function. It is possible that the terms negotiated from non-employment when $h_0 = 0$ are preferable to those negotiated when the outside offer is B and $h_0 = 1$.

- Rental rate is R_{iAB0} and worker has augmented human capital and an offer from B: The indeterminacy in this case is clearer and is also dependent on the utility function. In the separable/log case it does not matter because the income and substitution cancel out, so workers are indifferent between using the offer to renegotiate or ignoring it, but in other cases they will not be indifferent.

Since in both cases it is undetermined whether or not to use the offer from B to renegotiate the wage, we can not obtain simple expressions like the one above. We consider each of these two rental rates.

R_{iA00}: The solution depends on whether the worker wants to use an offer from B to renegotiate after human capital increases. I.e. there are two different cases depending on the whether

[5] For the others we know $R_{iAA0} = R_{iAA1} = \pi_{iA}$ and $R_{iBB0} = R_{iBB1} = \pi_{iB}$.

there is renegotiation.

Case 1: Workers prefers to renegotiate which happens when $[\beta V_{iA1}(\pi_{iA}) + (1-\beta)\, V_{iB1}(\pi_{iB})] \geq V_{iA1}\,(R_{iA00})$

Working through the algebra one can show

$$\begin{aligned}(\rho+\delta+\lambda_B^e+\lambda_A^e+\lambda_h)\,(1-\beta)\,V_{iA0}(\pi_{iA}) \\ =& u_{iA}\,(\pi_{iA}) - u_{iA}\,(R_{iA00}) + (1-\beta)\,[V_{iA0}(\pi_{iA}) - V_{iB0}(\pi_{iB})] \\ &+ \lambda_h \frac{u_{iA}\,(\pi_{iA}\psi_1) - u_{iA}\,(R_{iA00}\psi) + \lambda_B^e\,(1-\beta)\,[V_{iA1}(\pi_{iA}) + V_{iB1}(\pi_{iB})]}{\rho+\delta+\lambda_B^e+\lambda_A^e}.\end{aligned} \tag{E.3.8}$$

Case 2: Workers do not renegotiate which happens when $[\beta V_{iA1}(\pi_{iA}) + (1-\beta)\, V_{iB1}(\pi_{iB})] < V_{iA1}\,(R_{iA00})$.

We can solve the model to show

$$\begin{aligned}(\rho+\delta+\lambda_B^e+\lambda_A^e+\lambda_h)\,(1-\beta)\,V_{iA0}(\pi_{iA}) =& u_{iA}\,(\pi_{iA}) - u_{iA}\,(R_{iA00}) + (1-\beta)\,[V_{iA0}(\pi_{iA}) - V_{iB0}(\pi_{iB})] \\ &+ \lambda_h \frac{u_{iA}\,(\pi_{iA}\psi_1) - u_{iA}\,(R_{iA00}\psi)}{\rho+\delta+\lambda_A^e}.\end{aligned} \tag{E.3.9}$$

R_{iAB0}: Here again the solution depends on whether the worker wants to use an offer from B to renegotiate.

Case 1: Workers do not renegotiate: $V_{iA1}\,(R_{iAB0}) \geq \beta V_{iA1}(\pi_{iA}) + (1-\beta)\, V_{iB1}(\pi_{iB})$

$$(\rho+\delta+\lambda_h)\,(1-\beta)\,V_{iA0}(\pi_{iA}) = u_{iA}\,(\pi_{iA}) - u_{iA}\,(R_{iAB0}) + \lambda_h \frac{u_{iA}\,(\pi_{iA}\psi_1) - u_{iA}\,(R_{iAB0}\psi_1)}{\rho+\delta+\lambda_A^e}. \tag{E.3.10}$$

Case 2: Workers do renegotiate: $V_{iA1}\,(R_{iAB0}) < \beta V_{iA1}(\pi_{iA}) + (1-\beta)\, V_{iB1}(\pi_{iB})$

$$\begin{aligned}(\rho+\delta+\lambda_A^e+\lambda_h)\,(1-\beta)\,V_{iA0}(\pi_{iA}) =& u_{iA}\,(\pi_{iA}) - u_{iA}\,(R_{iAB0}) \\ &+\lambda_h \frac{u_{iA}\,(\pi_{iA}\psi_1) - u_{iA}\,(R_{iAB0}\psi) + \lambda_B^e\,(1-\beta)\,[V_{iA1}(\pi_{iA}) - V_{iB1}(\pi_{iB})]}{\rho+\delta+\lambda_B^e+\lambda_A^e}.\end{aligned} \tag{E.3.11}$$

Thus, for R_{iAB0} and R_{iA00} we have the four equations (E.3.8)-(E.3.11). While the expressions are more complicated, we get the same result. If $\beta = 1$ then $R_{iA00} = R_{iAB0} = \pi_{iA}$ with probability one and if $R_{iA00} = R_{iAB0} = \pi_{iA}$ then either $\beta = 1$ or workers are indifferent between all relevant options.

Second part when $C_i = A0$

The value β is irrelevant for the turnover decision and the parameters governing those decisions are identified, so we focus on the wage equations. It is very messy, but straight forward to derive the wages (formal derivation is available from the authors). After normalizing $U_{i00} = 0$, the two endogenous wages can be written as

$$\begin{aligned}
\varpi_{iA000} =& \Gamma_{0u}\left(\beta, \delta_i\right) U_{i01} + \Gamma_{0\pi}\left(\beta, \delta_i\right) \pi_{iA} + \Gamma_{i0v}\left(\beta, \delta_i\right) v_{iA} + \Gamma_{0\psi}\left(\beta, \delta_i\right) \log\left(\psi_1\right) \\
\varpi_{iA011} =& \Gamma_{1u}\left(\beta, \delta_i\right) U_{i01} + \Gamma_{1\pi}\left(\beta, \delta_i\right) \pi_{iA} + \Gamma_{i1v}\left(\beta, \delta_i\right) v_{iA} + \Gamma_{1\psi}\left(\beta, \delta_i\right) \log\left(\psi_1\right),
\end{aligned}$$

where the Γ are very messy terms which depend on parameters that we have shown are identified up to β. We have also identified ψ_1 and the distribution of π_{iA}.

To see that this model is not identified note that for any $\widetilde{\beta} \in [0,1)$, we can generically[6] find a value $\widetilde{U}_{i01}\left(\widetilde{\beta}\right)$ and $\widetilde{v}_{iA}\left(\widetilde{\beta}\right)$ that give the same wages. They are the solution to the two equations

$$\begin{aligned}
\varpi_{iA000} =& \Gamma_{0u}\left(\widetilde{\beta}, \delta_i\right) \widetilde{U}_{i01}\left(\widetilde{\beta}\right) + \Gamma_{0\pi}\left(\widetilde{\beta}, \delta_i\right) \pi_{iA} + \Gamma_{0v}\left(\widetilde{\beta}, \delta_i\right) \widetilde{v}_{iA}\left(\widetilde{\beta}\right) + \Gamma_{0\psi}\left(\beta, \delta_i\right) \log\left(\psi_1\right) \\
\varpi_{iA011} =& \Gamma_{1u}\left(\widetilde{\beta}, \delta_i\right) \widetilde{U}_{i01}\left(\widetilde{\beta}\right) + \Gamma_{1\pi}\left(\widetilde{\beta}, \delta_i\right) \pi_{iA} + \Gamma_{1v}\left(\widetilde{\beta}, \delta_i\right) \widetilde{v}_{iA}\left(\widetilde{\beta}\right) + \Gamma_{1\psi}\left(\beta, \delta_i\right) \log\left(\psi_1\right),
\end{aligned}$$

so as long as there is a solution to these equations, the model can not be identified.
The parameters take the form

$$\begin{aligned}
\Gamma_{0u}\left(\widetilde{\beta}, \delta_i\right) =& \frac{d_1\left(\delta_i\right)\left[\left(d_0\left(\delta_i\right)\beta - \lambda_A^e\right) a_{0u}\left(\beta, \delta_i\right) + \left(d_0\left(\delta_i\right)\left(1-\beta\right) - \delta_i\right) b_{0u}\left(\beta, \delta_i\right)\right] + \lambda_h\left[\lambda_A^e a_{1u}\left(\beta, \delta_i\right) + \delta_i c_{i1u}\left(\beta\right)\right]}{d_1\left(\delta_i\right) + \lambda_h} \\
\Gamma_{0\pi}\left(\widetilde{\beta}, \delta_i\right) =& \frac{d_1\left(\delta_i\right)\left[\left(d_0\left(\delta_i\right)\beta - \lambda_A^e\right) a_{0\pi}\left(\beta, \delta_i\right) + \left(d_0\left(\delta_i\right)\left(1-\beta\right) - \delta_i\right) b_{0\pi}\left(\beta, \delta_i\right)\right] + \lambda_h\left[\lambda_A^e a_{1\pi}\left(\beta, \delta_i\right) + \delta_i c_{i1\pi}\left(\beta\right)\right]}{d_1\left(\delta_i\right) + \lambda_h} \\
\Gamma_{0\pi}\left(\widetilde{\beta}, \delta_i\right) =& -1 + \frac{d_1\left(\delta_i\right)\left[\left(d_0\left(\delta_i\right)\beta - \lambda_A^e\right) a_{0\pi}\left(\beta, \delta_i\right) + \left(d_0\left(\delta_i\right)\left(1-\beta\right) - \delta_i\right) b_{0\pi}\left(\beta, \delta_i\right)\right] + \lambda_h\left[\lambda_A^e a_{1\pi}\left(\beta, \delta_i\right) + \delta_i c_{i1\pi}\left(\beta\right)\right]}{d_1\left(\delta_i\right) + \lambda_h} \\
\Gamma_{0\psi}\left(\beta, \delta_i\right) =& \frac{d_1\left(\delta_i\right)\left[\left(d_0\left(\delta_i\right)\beta - \lambda_A^e\right) a_{0\psi}\left(\beta, \delta_i\right) + \left(d_0\left(\delta_i\right)\left(1-\beta\right) - \delta_i\right) b_{0\psi}\left(\beta, \delta_i\right)\right] + \lambda_h\left[1 + \lambda_A^e a_{1\psi}\left(\beta, \delta_i\right) + \delta_i c_{i1\psi}\left(\beta\right)\right]}{d_1\left(\delta_i\right) + \lambda_h} \\
\Gamma_{1u}\left(\beta\right) =& -\delta_i c_{1u}\left(\beta, \delta_i\right) + \left[d_{i1}\beta - \lambda_A^e\right] a_{1u}\left(\beta, \delta_i\right) + d_{i1}\left(1-\beta\right) b_{1u}\left(\beta, \delta_i\right) \\
\Gamma_{1\pi}\left(\beta\right) =& -\delta_i c_{1\pi}\left(\beta, \delta_i\right) + \left[d_{i1}\beta - \lambda_A^e\right] a_{1\pi}\left(\beta, \delta_i\right) + d_{i1}\left(1-\beta\right) b_{1\pi}\left(\beta, \delta_i\right) \\
\Gamma_{1v}\left(\beta\right) =& -1 - \delta_i c_{1\pi}\left(\beta, \delta_i\right) + \left[d_{i1}\beta - \lambda_A^e\right] a_{1\psi}\left(\beta, \delta_i\right) + d_{i1}\left(1-\beta\right) b_{i1\pi}\left(\beta\right) \\
\Gamma_{1\psi}\left(\beta\right) =& -\delta_i c_{1\psi}\left(\beta, \delta_i\right) + \left[d_{i1}\beta - \lambda_A^e\right] a_{i1\psi}\left(\beta\right) + d_{i1}\left(1-\beta\right) b_{1\psi}\left(\beta, \delta_i\right),
\end{aligned}$$

[6] By generically we mean that as long as the two equations are not linearly dependent. There is no reason why they should be, so generically they will not but we can not rule out very special cases where they are linearly dependent.

with

$$
\begin{aligned}
\widetilde{a}_{1\pi}(\delta_i) =& \frac{1}{\rho+\delta_i} \\
\widetilde{a}_{1V}(\delta_i) =& \frac{\delta_i}{\rho+\delta_i} \\
\widetilde{b}_{1u}(\beta) =& \frac{1}{\rho+\beta\lambda_A^n} \\
\widetilde{b}_{1\pi}(\beta,\delta_i) =& \frac{\lambda_A^n\beta\widetilde{a}_{1\pi}(\delta_i)}{(\rho+\beta\lambda_A^n)} \\
\widetilde{b}_{1V}(\beta,\delta_i) =& \frac{\lambda_A^n\beta\widetilde{a}_{1V}(\delta_i)}{(\rho+\beta\lambda_A^n)} \\
P_A^* =& P^*\frac{\lambda_A^n}{\lambda_A^n+\lambda_A^n} \\
c_{1u}(\beta,\delta_i) =& \frac{(1-P_A^*\beta)\,\widetilde{b}_{1u}(\beta)}{1-P_A^*\beta\widetilde{a}_{1V}(\delta_i)-(1-P_A^*\beta)\,\widetilde{b}_{1V}(\beta,\delta_i)} \\
c_{1\pi}(\beta,\delta_i) = c_{1\psi}(\beta,\delta_i) =& \frac{P_A^*\beta\widetilde{a}_{i1\pi}+(1-P_A^*\beta)\,\widetilde{b}_{1\pi}(\beta,\delta_i)}{1-P_A^*\beta\widetilde{a}_{1V}(\delta_i)-(1-P_A^*\beta)\,\widetilde{b}_{1V}(\beta,\delta_i)} \\
b_{1u}(\beta,\delta_i) =& \widetilde{b}_{1u}(\beta)+\widetilde{b}_{1V}(\beta,\delta_i)\,c_{1u}(\beta,\delta_i) \\
b_{1\pi}(\beta,\delta_i) = b_{1\psi}(\beta,\delta_i) =& \widetilde{b}_{1\pi}(\beta,\delta_i)+\widetilde{b}_{1V}(\beta,\delta_i)\,c_{1\pi}(\beta,\delta_i) \\
a_{1u}(\beta,\delta_i) =& \widetilde{a}_{1V}(\delta_i)\,c_{1u}(\beta,\delta_i) \\
a_{1\pi}(\beta,\delta_i) = a_{1\psi}(\beta,\delta_i) =& \widetilde{a}_{1\pi}(\delta_i)+\widetilde{a}_{1V}(\delta_i)\,c_{1\pi}(\beta,\delta_i) \\
\widetilde{a}_{0\pi}(\delta_i) =& \frac{1}{\rho+\delta_i+\lambda_h} \\
\widetilde{a}_{0V}(\delta_i) =& \frac{\delta_i}{\rho+\delta_i+\lambda_h} \\
\widetilde{b}_{0\pi}(\beta,\delta_i) =& \frac{\lambda_A^n\beta\widetilde{a}_{0\pi}(\delta_i)}{(\rho+\beta\lambda_A^n)} \\
\widetilde{b}_{0V}(\beta,\delta_i) =& \frac{\lambda_A^n\beta\widetilde{a}_{0V}(\delta_i)}{(\rho+\beta\lambda_A^n)} \\
\widetilde{c}_0(\beta,\delta_i) =& \frac{P_A^*\beta\widetilde{a}_{0\pi}(\delta_i)+(1-P_A^*\beta)\,\widetilde{b}_{0\pi}(\beta,\delta_i)}{1-P_A^*\beta\widetilde{a}_{0V}(\delta_i)-(1-P_A^*\beta)\,\widetilde{b}_{0V}(\beta,\delta_i)} \\
c_{0u}(\beta,\delta_i) =& \widetilde{c}_0(\beta,\delta_i)\,\lambda_h a_{1u}(\beta,\delta_i) \\
c_{0\pi}(\beta,\delta_i) =& \widetilde{c}_0(\beta,\delta_i)\,[1+\lambda_h a_{1\pi}(\beta,\delta_i)] \\
c_{0\psi}(\beta,\delta_i) =& \widetilde{c}_0(\beta,\delta_i)\,\lambda_h a_{1\psi}(\beta,\delta_i) \\
b_{0u}(\beta,\delta_i) =& \left[\widetilde{b}_{0\pi}(\beta,\delta_i)+\widetilde{b}_{0V}(\beta,\delta_i)\,\widetilde{c}_0(\beta,\delta_i)\right]\lambda_h a_{1u}(\beta,\delta_i) \\
b_{0\pi}(\beta,\delta_i) =& \left[\widetilde{b}_{0\pi}(\beta,\delta_i)+\widetilde{b}_{0V}(\beta,\delta_i)\,\widetilde{c}_0(\beta,\delta_i)\right](1+\lambda_h a_{1\pi}(\beta,\delta_i)) \\
b_{0\psi}(\beta,\delta_i) =& \left[\widetilde{b}_{0\pi}(\beta,\delta_i)+\widetilde{b}_{0V}(\beta,\delta_i)\,\widetilde{c}_0(\beta,\delta_i)\right]\lambda_h a_{1\psi}(\beta,\delta_i) \\
a_{0u}(\beta,\delta_i) =& [\widetilde{a}_{0\pi}(\delta_i)+\widetilde{a}_{0V}(\delta_i)\,\widetilde{c}_0(\beta,\delta_i)]\,\lambda_h a_{1u}(\beta,\delta_i) \\
a_{0\pi}(\beta,\delta_i) =& [\widetilde{a}_{0\pi}(\delta_i)+\widetilde{a}_{0V}(\delta_i)\,\widetilde{c}_0(\beta,\delta_i)]\,(1+\lambda_h a_{1\pi}(\beta,\delta_i)) \\
a_{0\psi}(\beta,\delta_i) =& [\widetilde{a}_{0\pi}(\delta_i)+\widetilde{a}_{0V}(\delta_i)\,\widetilde{c}_0(\beta,\delta_i)]\,\lambda_h a_{1\psi}(\beta,\delta_i) \\
d_1(\delta_i) =& \rho+\delta_i+\lambda_A^e \\
d_0(\delta_i) =& \rho+\delta_i+\lambda_A^e+\lambda_h.
\end{aligned}
$$

However, there is a loose end. We also need to show that the new model with $\left(\widetilde{\beta},\widetilde{v}_{iA}\left(\widetilde{\beta}\right),\widetilde{U}_{i01}\left(\widetilde{\beta}\right)\right)$ produces the same choice ordering as the base model with (β, v_{iA}, U_{i01}). That is, even though the taste components are different, the worker would remain a $C_i = A0$ worker.

First, we show that this is the case for human capital equal to 1. To see this, we can write

$$\begin{aligned}(\rho+\delta+\lambda_A^e)\left[\beta V_{iA1}(\pi_{iA})+(1-\beta)\,V_{iA1}(\pi_{iA})\right]&=\log(\pi_{iA})+\log(\psi_1)+v_{iA}+\lambda_A^e V_{iA1}(\pi_{iA})+\delta_i V_{i01}^*\\(\rho+\delta+\lambda_A^e)\left[\beta V_{iA1}(\pi_{iA})+(1-\beta)\,V_{i01}\right]&=\varpi_{iA011}+v_{iA}+\lambda_A^e V_{iA1}(\pi_{iA})+\delta_i V_{i01}^*,\end{aligned}$$

so

$$(\rho+\delta+\lambda_A^e)\,(1-\beta)\left[V_{iA1}(\pi_{iA})-V_{i01}\right]=\log(\pi_{iA})+\log(\psi_1)-\varpi_{iA011},$$

so $V_{iA1}(\pi_{iA})>V_{i01}$ implies $\log(\pi_{iA})+\log(\psi_1)>\varpi_{iA011}$.

We can use the same argument in reverse to get the following result. That is define $\widetilde{V}_{iA1}\left(\pi_{iA};\widetilde{\beta}\right)$ and $\widetilde{V}_{i01}\left(\widetilde{\beta}\right)$ to be the value functions implied by the alternative model, since this is just a different parameterization of the same model, it must be the case that

$$(\rho+\delta+\lambda_A^e)\left(1-\widetilde{\beta}\right)\left[\widetilde{V}_{iA1}\left(\pi_{iA};\widetilde{\beta}\right)-\widetilde{V}_{i01}\left(\widetilde{\beta}\right)\right]=\log(\pi_{iA})+\log(\psi_1)-\varpi_{iA011},$$

so $\log(\pi_{iA})+\log(\psi_1)>\varpi_{iA011}$ implies $\widetilde{V}_{iA1}\left(\pi_{iA};\widetilde{\beta}\right)>\widetilde{V}_{i01}\left(\widetilde{\beta}\right)$.

Now consider the case without augmented human capital. We can solve the model to show

$$(\rho+\delta_i+\lambda_A^e+\lambda_h)\,(1-\beta)\left[V_{iA0}(\pi_{iA})-V_{i00}(\beta)\right]=\frac{\rho+\delta_i+\lambda_A^e+\lambda_h}{\rho+\delta_i+\lambda_A^e}\left[\log(\pi_{iA})-\varpi_{iA000}\right].$$

Thus, $V_{iA0}(\pi_{iA})>V_{i00}(\beta)$ implies $\log(\pi_{iA})>\varpi_{iA000}$.

Again, we show the inverse. The new model is an alternative parameterization of the old one so we can write

$$(\rho+\delta_i+\lambda_A^e+\lambda_h)\left(1-\widetilde{\beta}\right)\left[\widetilde{V}_{iA0}\left(\pi_{iA};\widetilde{\beta}\right)-\widetilde{V}_{i00}\left(\widetilde{\beta}\right)\right]=\frac{\rho+\delta_i+\lambda_A^e+\lambda_h}{\rho+\delta_i+\lambda_A^e}\left[\log(\pi_{iA})-\varpi_{iA000}\right],$$

and thus $\widetilde{V}_{iA0}\left(\pi_{iA};\widetilde{\beta}\right)>\widetilde{V}_{i00}\left(\widetilde{\beta}\right)$ and the choice ordering is the same in the case without augmented human capital.

This completes the second part of the proof for the $C_i=A0$ case. Substituting A with B gives the proof for the $C_i=B0$ case.

Second part when $C_i = AB$

In this case there are 6 endogenous wages $\varpi_{iA000}, \varpi_{iAB00}, \varpi_{iA011}, \varpi_{iAB11}, \varpi_{iB000}$, and ϖ_{iB011}. Working through the model (details available on request from the authors) one can show

$$\varpi_{iAB00} = \log(\pi_{iA}) - (\rho + \delta_i + \lambda_A^e)(1-\beta)\left[\frac{\Delta_i}{\rho + \delta_i + \beta\lambda_A^e}\right] \tag{E.3.12}$$

$$= \log(\pi_{iA}) + \Gamma_{AB0\Delta}(\beta, \delta_i)\,\Delta_i \tag{E.3.13}$$

$$\varpi_{iAB11} = \varpi_{iAB00} + \log(\psi_1) \tag{E.3.14}$$

$$\varpi_{iA011} = \Gamma_{A01u}(\beta, \delta_i)\,U_{i01} + \Gamma_{A01\pi}(\beta, \delta_i)\log(\pi_{iA}) + \Gamma_{A01v}(\beta, \delta_i)\,v_{iA} + \Gamma_{A01\psi}(\beta, \delta_i)\log(\psi_1) + \Gamma_{A01\Delta}(\beta, \delta_i)\,\Delta_i \tag{E.3.15}$$

$$\varpi_{iA000} = \Gamma_{A00u}(\beta, \delta_i)\,U_{i01} + \Gamma_{A00\pi}(\beta, \delta_i)\log(\pi_{iA}) + \Gamma_{A00v}(\beta, \delta_i)\,v_{iA} + \Gamma_{A00\psi}(\beta, \delta_i)\log(\psi_1) + \Gamma_{A00\Delta}(\beta, \delta_i)\,\Delta_i \tag{E.3.16}$$

$$\varpi_{iB011} = \log(\pi_{iB}) - \varpi_{iAB11} + \varpi_{iA011} + \log(\psi_1) \tag{E.3.17}$$

$$\varpi_{iB000} = \log(\pi_{iB}) - \varpi_{iAB00} + \varpi_{iA000} + \log(\pi_{iB}) - \varpi_{iAB00} + \varpi_{iA000}, \tag{E.3.18}$$

where

$$
\begin{aligned}
\Gamma_{AB0\Delta}(\beta,\delta_i) =& \frac{\left(\rho+\delta_i+\lambda_A^e\right)(1-\beta)}{\rho+\delta_i+\beta\lambda_A^e} \\
\Gamma_{A01u}(\beta,\delta_i) =& -\delta_i c_{1u}(\beta,\delta_i) + \left[d_1(\delta_i)\beta-\Lambda^e\right] a_{1u}(\beta,\delta_i) + d_1(\delta_i)(1-\beta)\, b_{1u}(\beta,\delta_i) \\
\Gamma_{A01\pi}(\beta,\delta_i) =& -\delta_i c_{1\pi}(\beta,\delta_i) + \left[d_1(\delta_i)\beta-\Lambda^e\right] a_{1\pi}(\beta) + d_1(\delta_i)(1-\beta)\, b_{1\pi}(\beta,\delta_i) \\
\Gamma_{A01v}(\beta,\delta_i) =& -1-\delta_i c_{1\pi}(\beta,\delta_i) + \left[d_1(\delta_i)\beta-\Lambda^e\right] a_{1\pi}(\beta) + d_1(\delta_i)(1-\beta)\, b_{1\pi}(\beta,\delta_i) \\
\Gamma_{A01\psi}(\beta,\delta_i) =& -\delta_i c_{1\psi}(\beta,\delta_i) + \left[d_1(\delta_i)\beta-\Lambda^e\right] a_{1\psi}(\beta,\delta_i) + d_1(\delta_i)(1-\beta)\, b_{1\psi}(\beta,\delta_i) \\
\Gamma_{A01\Delta}(\beta,\delta_i) =& \delta_i c_{1\Delta}(\beta,\delta_i) - \left[d_1(\delta_i)\beta-\Lambda^e\right] a_{1\psi}(\beta) - d_1(\delta_i)(1-\beta)\, b_{1\Delta}(\beta,\delta_i) + \lambda_B^e(1-\beta)\,\widetilde{a}_{1\Delta}(\delta_i) \\
\Gamma_{A01u}(\beta,\delta_i) =& -\delta_i c_{1u}(\beta,\delta_i) + \left[d_1(\delta_i)\beta-\Lambda^e\right] a_{1u}(\beta,\delta_i) + d_1(\delta_i)(1-\beta)\, b_{1u}(\beta,\delta_i) \\
\Gamma_{A01\pi}(\beta,\delta_i) =& -\delta_i c_{1\pi}(\beta,\delta_i) + \left[d_1(\delta_i)\beta-\Lambda^e\right] a_{1\pi}(\beta) + d_1(\delta_i)(1-\beta)\, b_{1\pi}(\beta,\delta_i) \\
\Gamma_{A01v}(\beta,\delta_i) =& -1-\delta_i c_{1\pi}(\beta,\delta_i) + \left[d_1(\delta_i)\beta-\Lambda^e\right] a_{1\pi}(\beta) + d_1(\delta_i)(1-\beta)\, b_{1\pi}(\beta,\delta_i) \\
\Gamma_{A01\psi}(\beta,\delta_i) =& -\delta_i c_{1\psi}(\beta,\delta_i) + \left[d_1(\delta_i)\beta-\Lambda^e\right] a_{1\psi}(\beta,\delta_i) + d_1(\delta_i)(1-\beta)\, b_{1\psi}(\beta,\delta_i) \\
\Gamma_{A01\Delta}(\beta,\delta_i) =& \delta_i c_{1\Delta}(\beta,\delta_i) - \left[d_1(\delta_i)\beta-\Lambda^e\right] a_{1\psi}(\beta) - d_1(\delta_i)(1-\beta)\, b_{1\Delta}(\beta,\delta_i) + \lambda_B^e(1-\beta)\,\widetilde{a}_{1\Delta}(\delta_i) \\
\Gamma_{A00u}(\beta,\delta_i) =& \frac{d_1(\delta_i)\left[\left(d_0(\delta_i)\beta-\Lambda^e\right)a_{0u}(\beta,\delta_i) + \left(d_0(\delta_i)(1-\beta)\right)b_{0u}(\beta,\delta_i) - \delta_i c_{0u}(\beta,\delta_i) - \lambda_h\left[\Lambda_A^e a_{1u}(\beta,\delta_i)+\delta_i c_{1u}(\beta,\delta_i)\right]\right]}{d_1(\delta_i)+\lambda_h} \\
\Gamma_{A00\pi}(\beta,\delta_i) =& \frac{d_1(\delta_i)\left[\left(d_0(\delta_i)\beta-\Lambda^e\right)a_{0\pi}(\beta,\delta_i) + \left(d_0(\delta_i)(1-\beta)\right)b_{0\pi}(\beta,\delta_i) - \delta_i c_{0\pi}(\beta,\delta_i) - \lambda_h\left[\Lambda_A^e a_{1\pi}(\beta,\delta_i)+\delta_i c_{1\pi}(\beta,\delta_i)\right]\right]}{d_1(\delta_i)+\lambda_h} \\
\Gamma_{A00v}(\beta,\delta_i) =& -1 + \frac{d_1(\delta_i)\left[\left(d_0(\delta_i)\beta-\Lambda^e\right)a_{0\pi}(\beta,\delta_i) + \left(d_0(\delta_i)(1-\beta)\right)b_{0\pi}(\beta,\delta_i) - \delta_i c_{0\pi}(\beta,\delta_i) - \lambda_h\left[\Lambda_A^e a_{1\pi}(\beta,\delta_i)+\delta_i c_{1\pi}(\beta,\delta_i)\right]\right]}{d_1(\delta_i)+\lambda_h} \\
\Gamma_{A00\psi}(\beta,\delta_i) =& \frac{d_1(\delta_i)\left[\left(d_0(\delta_i)\beta-\Lambda^e\right)a_{0\psi}(\beta,\delta_i) + \left(d_0(\delta_i)(1-\beta)\right)b_{0\psi}(\beta,\delta_i) - \delta_i c_{0\pi}(\beta,\delta_i) - \lambda_h\left[1+\Lambda_A^e a_{1\psi}(\beta,\delta_i)+\delta_i c_{1\psi}(\beta,\delta_i)\right]\right]}{d_1(\delta_i)+\lambda_h} \\
\Gamma_{A00\Delta}(\beta,\delta_i) =& -\frac{d_1(\delta_i)\left[\left(d_0(\delta_i)\beta-\Lambda^e\right)a_{0\Delta}(\beta,\delta_i) + \left(d_0(\delta_i)(1-\beta)\right)b_{0\Delta}(\beta,\delta_i) - \delta_i c_{0\Delta}(\beta,\delta_i) - \lambda_h\left[\lambda_B^e(1-\beta)+\Lambda_A^e a_{1\Delta}(\beta,\delta_i)+\delta_i c_{1\Delta}(\beta,\delta_i)\right]\right]}{d_1(\delta_i)+\lambda_h} \\
\Delta_i =& \log(\pi_{iA}) + v_{iA} - \log(\pi_{iB}) - v_{iB}.
\end{aligned}
$$

and

$$
\begin{aligned}
\widetilde{a}_{1\pi}(\delta_i) =& \frac{1}{\rho+\delta_i}, \widetilde{a}_{1V}(\delta_i) = \frac{\delta_i}{\rho+\delta_i}, \widetilde{a}_{1\Delta}(\delta_i) = \frac{1}{\rho+\delta_i+\beta\lambda_A^e} \\
\widetilde{b}_{10}(\beta) =& \frac{1}{\rho+\beta\Lambda^n}, \widetilde{b}_{1\pi}(\beta,\delta_i) = \frac{\Lambda^n\beta\widetilde{a}_{1\pi}(\delta_i)}{\rho+\beta\Lambda^n}, \widetilde{b}_{1V}(\beta,\delta_i) = \frac{\Lambda^n\beta\widetilde{a}_{1V}(\delta_i)}{\rho+\beta\Lambda^n}, \widetilde{b}_{1\Delta}(\beta,\delta_i) = \frac{\lambda_B^n\beta\widetilde{a}_{1\Delta}(\delta_i)}{\rho+\beta\Lambda^n} \\
c_{1u}(\beta,\delta_i) =& \frac{(1-\beta P^*)\widetilde{b}_{10}(\beta)}{1-P^*\beta\widetilde{a}_{1V}(\delta_i)-(1-\beta P^*)\widetilde{b}_{1V}(\beta,\delta_i)(\beta)} \\
c_{1\pi}(\beta,\delta_i) = c_{1\psi}(\beta,\delta_i) =& \frac{P^*\beta\widetilde{a}_{1\pi}(\delta_i)+(1-\beta P^*)\widetilde{b}_{1\pi}(\beta)}{1-P^*\beta\widetilde{a}_{1V}(\delta_i)-(1-\beta P^*)\widetilde{b}_{1V}(\beta,\delta_i)} \\
c_{i1\Delta}(\beta,\delta_i) =& \frac{\frac{\lambda_B^n}{\Lambda^n}\beta\widetilde{a}_{1\Delta}(\delta_i)+(1-\beta P^*)\widetilde{b}_{1\Delta}(\beta,\delta_i)}{1-P^*\beta\widetilde{a}_{1V}(\delta_i)-(1-\beta P^*)\widetilde{b}_{1V}(\beta,\delta_i)} \\
b_{1u}(\beta,\delta_i) =& \widetilde{b}_{1u}(\beta)+\widetilde{b}_{1V}(\beta,\delta_i)c_{1u}(\beta,\delta_i) \\
b_{1\pi}(\beta,\delta_i) = b_{1\psi}(\beta,\delta_i) =& \widetilde{b}_{1\pi}(\beta,\delta_i)+\widetilde{b}_{1V}(\beta,\delta_i)c_{1\pi}(\beta,\delta_i) \\
b_{1\Delta}(\beta,\delta_i) =& \widetilde{b}_{1\Delta}(\beta,\delta_i)+\widetilde{b}_{1V}(\beta,\delta_i)c_{1\Delta}(\beta,\delta_i) \\
a_{1u}(\beta,\delta_i) =& \widetilde{a}_{i12}c_{i10}(\beta) \\
a_{1\pi}(\beta) = a_{1\psi}(\beta) =& \widetilde{a}_{1\pi}(\delta_i)+\widetilde{a}_{1V}(\delta_i)c_{1\pi}(\beta,\delta_i) \\
a_{1\Delta}(\beta) =& \widetilde{a}_{1V}(\delta_i)c_{1\Delta}(\beta,\delta_i) \\
\widetilde{a}_{0\pi}(\delta_i) =& \frac{1}{\rho+\delta_i+\lambda_h}, \widetilde{a}_{0V}(\delta_i) = \frac{\delta_i}{\rho+\delta_i+\lambda_h}, \widetilde{a}_{0\Delta}(\delta_i) = \frac{1}{\rho+\delta_i+\beta\lambda_A^e} \\
\widetilde{b}_{0\pi}(\beta,\delta_i) =& \frac{\Lambda^n\beta\widetilde{a}_{0\pi}(\delta_i)}{\rho+\beta\Lambda^n}, \widetilde{b}_{0V}(\beta,\delta_i) = \frac{\Lambda^n\beta\widetilde{a}_{0V}(\delta_i)}{\rho+\beta\Lambda^n}, \widetilde{b}_{0\Delta}(\beta,\delta_i) = \frac{\lambda_B^n\beta\widetilde{a}_{0\Delta}(\delta_i)}{\rho+\beta\Lambda^n} \\
\widetilde{c}_{0\pi}(\beta,\delta_i) =& \frac{P^*\beta\widetilde{a}_{0\pi}(\delta_i)+(1-\beta P^*)\widetilde{b}_{0\pi}(\beta,\delta_i)}{1-P^*\beta\widetilde{a}_{0\Delta}(\delta_i)-(1-\beta P^*)\widetilde{b}_{0V}(\beta,\delta_i)} \\
\widetilde{c}_{0\Delta}(\beta,\delta_i) =& \frac{\frac{\lambda_B^n}{\Lambda^n}\beta\widetilde{a}_{0\Delta}(\delta_i)+(1-\beta P^*)\widetilde{b}_{0\Delta}(\beta,\delta_i)}{1-P^*\beta\widetilde{a}_{0\Delta}(\delta_i)-(1-\beta P^*)\widetilde{b}_{0V}(\beta,\delta_i)(\beta)} \\
c_{0u}(\beta,\delta_i) =& \widetilde{c}_{0\pi}(\beta,\delta_i)\lambda_h a_{1u}(\beta,\delta_i) \\
c_{0\pi}(\beta,\delta_i) =& \widetilde{c}_{0\pi}(\beta,\delta_i)(1+\lambda_h a_{1\pi}(\beta)) \\
c_{0\psi}(\beta,\delta_i) =& \widetilde{c}_{0\pi}(\beta,\delta_i)\lambda_h a_{1\psi}(\beta) \\
c_{0\Delta}(\beta,\delta_i) =& \widetilde{c}_{0\pi}(\beta,\delta_i)\lambda_h a_{1\Delta}(\beta)+\widetilde{c}_{0\Delta}(\beta,\delta_i) \\
b_{0u}(\beta,\delta_i) =& \left[\widetilde{b}_{0\pi}(\beta,\delta_i)+\widetilde{b}_{0V}(\beta,\delta_i)\widetilde{c}_{0\pi}(\beta,\delta_i)\right]\lambda_h a_{1u}(\beta,\delta_i) \\
b_{0\pi}(\beta,\delta_i) =& \left[\widetilde{b}_{i01}(\beta)+\widetilde{b}_{i02}(\beta)\widetilde{c}_{i01}(\beta)\right](1+\lambda_h a_{1\pi}(\beta)) \\
b_{0\psi}(\beta,\delta_i) =& \left[\widetilde{b}_{i01}(\beta)+\widetilde{b}_{i02}(\beta)\widetilde{c}_{i01}(\beta)\right]\lambda_h a_{1\psi}(\beta) \\
b_{0\Delta}(\beta,\delta_i) =& \left[\widetilde{b}_{i01}(\beta)+\widetilde{b}_{i02}(\beta)\widetilde{c}_{i01}(\beta)\right]\lambda_h a_{1\Delta}(\beta)+\widetilde{b}_{0\Delta}(\beta,\delta_i)+\widetilde{b}_{0V}(\beta,\delta_i)\widetilde{c}_{0\Delta}(\beta,\delta_i) \\
a_{0u}(\beta,\delta_i) =& \left[\widetilde{a}_{0\pi}(\beta,\delta_i)+\widetilde{a}_{0V}(\beta,\delta_i)\widetilde{c}_{0\pi}(\beta,\delta_i)\right]\lambda_h a_{1u}(\beta,\delta_i) \\
a_{0\pi}(\beta,\delta_i) =& \left[\widetilde{a}_{0\pi}(\beta,\delta_i)+\widetilde{a}_{0V}(\beta,\delta_i)\widetilde{c}_{0\pi}(\beta,\delta_i)\right](1+\lambda_h a_{1\pi}(\beta)) \\
a_{0\psi}(\beta,\delta_i) =& \left[\widetilde{a}_{0\pi}(\beta,\delta_i)+\widetilde{a}_{0V}(\beta,\delta_i)\widetilde{c}_{0\pi}(\beta,\delta_i)\right]\lambda_h a_{1\psi}(\beta) \\
a_{0\Delta}(\beta,\delta_i) =& \left[\widetilde{a}_{0\pi}(\beta,\delta_i)+\widetilde{a}_{0V}(\beta,\delta_i)\widetilde{c}_{0\pi}(\beta,\delta_i)\right]\lambda_h a_{1\Delta}(\beta)+\widetilde{b}_{0\Delta}(\beta,\delta_i)+\widetilde{b}_{0V}(\beta,\delta_i)\widetilde{c}_{0\Delta}(\beta,\delta_i) \\
d_1(\delta_i) =& (\rho+\delta_i+\Lambda^e).
\end{aligned}
$$

To show the model is not identified note that for any $\widetilde{\beta} \in [0,1)$, we can generically find values $\widetilde{U}_{i01}\left(\widetilde{\beta}\right)$, $\widetilde{v}_{iA}\left(\widetilde{\beta}\right)$, and $\widetilde{v}_{iB}\left(\widetilde{\beta}\right)$ that give the same wages. They are the solution to the linear equations

$$
\begin{aligned}
\varpi_{iAB00} =& \log\left(\pi_{iA}\right) + \Gamma_{AB0\Delta}\left(\widetilde{\beta}, \delta_i\right)\Delta_i & \text{(E.3.19)}\\
\varpi_{iA011} =& \Gamma_{A01u}\left(\widetilde{\beta}, \delta_i\right)\widetilde{U}_{i01}\left(\widetilde{\beta}\right) + \Gamma_{A01\pi}\left(\widetilde{\beta}, \delta_i\right)\log\left(\pi_{iA}\right) + \Gamma_{A01v}\left(\widetilde{\beta}, \delta_i\right)\widetilde{v}_{iA}\left(\widetilde{\beta}\right) & \\
& + \Gamma_{A01\psi}\left(\widetilde{\beta}, \delta_i\right)\log(\psi_1) + \Gamma_{A01\Delta}\left(\widetilde{\beta}, \delta_i\right)\Delta_i & \text{(E.3.20)}\\
\varpi_{iA000} =& \Gamma_{A00u}\left(\widetilde{\beta}, \delta_i\right)\widetilde{U}_{i01}\left(\widetilde{\beta}\right) + \Gamma_{A00\pi}\left(\widetilde{\beta}, \delta_i\right)\log\left(\pi_{iA}\right) + \Gamma_{A00v}\left(\widetilde{\beta}, \delta_i\right)\widetilde{v}_{iA}\left(\widetilde{\beta}\right) & \\
& + \Gamma_{A00\psi}\left(\widetilde{\beta}, \delta_i\right)\log(\psi_1) + \Gamma_{A00\Delta}\left(\widetilde{\beta}, \delta_i\right)\Delta_i & \text{(E.3.21)}\\
\widetilde{\Delta}_i =& \log(\pi_{iA}) + \widetilde{v}_{iA}\left(\widetilde{\beta}\right) - \log(\pi_{iB}) - \widetilde{v}_{iB}\left(\widetilde{\beta}\right). & \text{(E.3.22)}
\end{aligned}
$$

The other three wages are still determined by (E.3.14), (E.3.18), and (E.3.18), which do not directly include β or values of the non-pecuniary benefits (i.e. the v's).

This shows that the model produces the same wages. We also need to show it produces the same choice ordering. That is, with the different non-pecuniary benefits (i.e. $\widetilde{v}$'s), we need to show that the worker would make the same choices.

As above define $\widetilde{V}_{iA0}\left(\pi_{iA}; \widetilde{\beta}\right)$, $\widetilde{V}_{iB0}\left(\pi_{iB}; \widetilde{\beta}\right)$, $\widetilde{V}_{i00}\left(\widetilde{\beta}\right)$, $\widetilde{V}_{iA1}\left(\pi_{iA}; \widetilde{\beta}\right)$, $\widetilde{V}_{iB1}\left(\pi_{iB}; \widetilde{\beta}\right)$ and $\widetilde{V}_{i01}\left(\widetilde{\beta}\right)$ to be the value functions implied by the alternative model, with all parameters remaining the same except $\widetilde{\beta}$, $\widetilde{U}_{i01}\left(\widetilde{\beta}\right)$, $\widetilde{v}_{iA}\left(\widetilde{\beta}\right)$, and $\widetilde{v}_{iB}\left(\widetilde{\beta}\right)$. We need to show that

$$
\begin{aligned}
\widetilde{V}_{i00}\left(\widetilde{\beta}\right) \leq& \widetilde{V}_{iB0}\left(\pi_{iB}; \widetilde{\beta}\right) \leq \widetilde{V}_{iA0}\left(\pi_{iA}; \widetilde{\beta}\right)\\
\widetilde{V}_{i01}\left(\widetilde{\beta}\right) \leq& \widetilde{V}_{iB1}\left(\pi_{iB}; \widetilde{\beta}\right) \leq \widetilde{V}_{iA1}\left(\pi_{iA}; \widetilde{\beta}\right).
\end{aligned}
$$

One result of the model is that

$$
V_{iA0}\left(\pi_{iA}\right) - V_{iB0}\left(\pi_{iB}\right) = \frac{\Delta_i}{\rho + \delta_i + \beta\lambda_A^e},
$$

which implies that $\Delta_i = \log(\pi_{iA}) + v_{iA} - \log(\pi_{iB}) - v_{iB} \geq 0$. Combined with (E.3.12) this implies that $\varpi_{iAB00} \leq \log\left(\pi_{iA}\right)$. When we plug this into (E.3.19) it implies that $\widetilde{\Delta}_i = \pi_{iA} + \widetilde{v}_{iA}\left(\widetilde{\beta}\right) - \pi_{iB} - \widetilde{v}_{iB}\left(\widetilde{\beta}\right) \geq 0$ and the model also implies that

$$
\widetilde{V}_{iA0}\left(\pi_{iA}; \widetilde{\beta}\right) - \widetilde{V}_{iB0}\left(\pi_{iB}; \widetilde{\beta}\right) = \frac{\widetilde{\Delta}_i}{\rho + \delta_i + \widetilde{\beta}\lambda_A^e},
$$

$$
\widetilde{V}_{iA1}\left(\pi_{iA}; \widetilde{\beta}\right) - \widetilde{V}_{iB1}\left(\pi_{iB}; \widetilde{\beta}\right) = \frac{\widetilde{\Delta}_i}{\rho + \delta_i + \widetilde{\beta}\lambda_A^e},
$$

and thus $\widetilde{V}_{iA0}\left(\pi_{iA}; \widetilde{\beta}\right) \geq \widetilde{V}_{iB0}\left(\pi_{iB}; \widetilde{\beta}\right)$ and $\widetilde{V}_{iA1}\left(\pi_{iA}; \widetilde{\beta}\right) \geq \widetilde{V}_{iB1}\left(\pi_{iB}; \widetilde{\beta}\right)$.

To show that $\widetilde{V}_{i01}\left(\widetilde{\beta}\right) \leq \widetilde{V}_{iB1}\left(\pi_{iB}; \widetilde{\beta}\right)$, we use the fact that the model implies that

$$
\left(\rho + \delta + \lambda_A^e + \lambda_B^e\right)(1 - \beta)\left[V_{iB1}(\pi_{iB}) - V_{i01}\right] = \varpi_{iAB11} - \varpi_{iA011}.
$$

Thus, $V_{iB1}(\pi_{iB}) \geq V_{i01}$ implies $\varpi_{iAB11} \geq \varpi_{iA011}$. But then the version of this equation with $\widetilde{\beta}$ is

$$(\rho + \delta + \lambda_A^e + \lambda_B^e)\left(1 - \widetilde{\beta}\right)\left[\widetilde{V}_{iB1}\left(\pi_{iB}; \widetilde{\beta}\right) - \widetilde{V}_{i01}\left(\widetilde{\beta}\right)\right] = \varpi_{iAB11} - \varpi_{iA011},$$

so $\varpi_{iAB11} \geq \varpi_{iA011}$ implies $\widetilde{V}_{iB1}\left(\pi_{iB}; \widetilde{\beta}\right) \geq \widetilde{V}_{i01}\left(\widetilde{\beta}\right)$.

We use an analogous argument to show that $\widetilde{V}_{i00}\left(\widetilde{\beta}\right) \leq \widetilde{V}_{iB0}\left(\pi_{iB}; \widetilde{\beta}\right)$. We can show that

$$\begin{aligned}(\rho + \delta + \lambda_B^e + \lambda_A^e + \lambda_h)(1 - \beta)\left[V_{iB0}(\pi_{iB}) - V_{i00}\right] = &\varpi_{iAB00} - \varpi_{iA000} \\ &+ \lambda_h \frac{\varpi_{iAB00} - \varpi_{iA000}}{\rho + \delta + \lambda_B^e + \lambda_A^e},\end{aligned}$$

so $V_{iB0}(\pi_{iB}) \geq V_{i00}$ implies $\varpi_{iAB00} \geq \varpi_{iA000}$. We can then write the analogue of this expression with the $\widetilde{\beta}$ alternative

$$\begin{aligned}(\rho + \delta + \lambda_B^e + \lambda_A^e + \lambda_h)\left(1 - \widetilde{\beta}\right)\left[\widetilde{V}_{iB0}\left(\pi_{iB}; \widetilde{\beta}\right) - \widetilde{V}_{i00}\left(\widetilde{\beta}\right)\right] = &\varpi_{iAB00} - \varpi_{iA000} \\ &+ \lambda_h \frac{\varpi_{iAB00} - \varpi_{iA000}}{\rho + \delta + \lambda_B^e + \lambda_A^e},\end{aligned}$$

which implies $\widetilde{V}_{iB0}\left(\pi_{iB}; \widetilde{\beta}\right) \geq \widetilde{V}_{i00}\left(\widetilde{\beta}\right)$. Thus, we have proven that

$$\begin{aligned}\widetilde{V}_{i00}\left(\widetilde{\beta}\right) \leq &\widetilde{V}_{iB0}\left(\pi_{iB}; \widetilde{\beta}\right) \leq \widetilde{V}_{iA0}\left(\pi_{iA}; \widetilde{\beta}\right) \\ \widetilde{V}_{i01}\left(\widetilde{\beta}\right) \leq &\widetilde{V}_{iB1}\left(\pi_{iB}; \widetilde{\beta}\right) \leq \widetilde{V}_{iA1}\left(\pi_{iA}; \widetilde{\beta}\right),\end{aligned}$$

and thereby both wages and choices are the same and thus the model with β , U_{i01}, v_{iA} , and v_{iB} can not be distinguished from the model with $\widetilde{\beta}$, $\widetilde{U}_{i01}\left(\widetilde{\beta}\right)$, $\widetilde{v}_{iA}\left(\widetilde{\beta}\right)$, and $\widetilde{v}_{iB}\left(\widetilde{\beta}\right)$.

D.5 Theorem D.1

Assumption D.2 *The econometrician observes the full history of job type spells with start and stop dates as well as the value of j at each job. The econometrician does not record job switches within job type.*

Theorem D.1 *Under Assumptions 1, D.2, and 4 with the data generated by the model exposited in the general model section , we can identify $\lambda_A^n, \lambda_B^n, P^*$, the distribution of C_i, and the distribution of δ_i conditional on C_i over the support of C_i for which $C_i \neq 0$. If $Pr(C_i = AB) > 0$ we can identify λ_A^e and if $Pr(C_i = BA) > 0$ we can identify λ_B^e.*

Proof

This is very similar to the proof of Theorem 1. We leave this as largely self contained, so it repeats many of the arguments we make in that proof.

We start by showing that we can identify $\lambda_A^n, \lambda_B^n, \lambda_A^e, \lambda_B^e, P^*$, the sample probabilities of C_i, and the distribution of δ_i (denote it F_δ) without using data on wages. A major complication is P^* because when we observe a job-to job-transition, we do not know whether it was voluntary or involuntary.

We will use $P(c)$ as shorthand notation for $Pr(C_i = c)$ with $c \in (0, B0, A0, BA, AB)$.

Identification of λ_A^n and λ_B^n

This is easier than in the base case. Since we observe workers forever, we know the ones who would accept both A and B jobs from non-employment (because they will eventually work for both). Condition on $C_i \in \{AB, BA\}$. The probability that the first firm is a B type firm is

$$P_B \equiv \frac{\lambda_B^n}{\lambda_A^n + \lambda_B^n}.$$

We define P_A in an analogous manner.

Continue to condition on $C_i \in \{AB, BA\}$. The hazard rate to the first job is $\lambda_A^n + \lambda_B^n$, so it is identified. From $\lambda_A^n + \lambda_B^n$ and P_B, we can identify λ_A^n and λ_B^n.

Identification of P^*, $P(AB|AB, BA)$, λ_A^e, λ_B^e, and the distribution of δ_i conditional on $C_i = AB$ and on $C_i = BA$

For this part of the proof, we will make use of three different employment spells. We can condition on individuals whose first three spells satisfy these conditions. Note that by employment spells we mean that there is a period of non-employment between them.

- The first begins at an A type firm and we follow it until the firm (type) spell ends. This can end with a job to job move to a B or with a non-employment spell. Let v_{1i} be the hazard rate of this spell ending (through either channel) for individual i, and let T_{1i} be the duration of this spell.

- The second begins at an B type firm and we follow it until the firm (type) spell ends. As above this can end with either a job to job move or to non-employment spell. Let v_{2i} be the hazard rate of this spell ending for individual i, and let T_{2i} be the duration of this spell.

- The third can begin at either type of firm and we follow the employment spell until it ends at non-employment. Let v_{3i} be the hazard rate to non-employment.

To be in this sample, the worker must be willing to take both an A type job and a B type job, so either $C_i = AB$ or $C_i = BA$. Since $P(AB) + P(BA) > 0$ we know that the sequence above can be observed in the data.

From data on the joint duration we can estimate the joint survivor function

$$Pr\left(T_{1i} \geq t_1, T_{2i} \geq t_2, T_{3i} \geq t_3\right) = \int e^{-v_{1i}t_1 - v_{2i}t_2 - v_{3i}t_3} dG\left(v_i\right),$$

where G is the conditional distribution of $v_i \equiv \left(v_{1i}, v_{2i}, v_{3i}\right)$. Note that this is the Laplace transform of G and one can invert the Laplace transform to identify G.

This is a random sample of BA and AB types because the two groups receive offers from A and B at the same rate. Thus in this sample

$$P\left(AB|AB, BA\right) \equiv \frac{P(AB)}{P(AB) + P(AB)}.$$

When $C_i = AB$,

$$v_{1i} = \delta_i\left[1 - P^* P_A\right], v_{2i} = \delta_i\left[1 - P^* P_B\right] + \lambda_A^e, v_{3i} = \delta_i\left[1 - P^*\right],$$

and when $C_i = BA$

$$v_{1i} = \delta_i\left[1 - P^* P_A\right] + \lambda_B^e, v_{2i} = \delta_i\left[1 - P^* P_B\right], v_{3i} = \delta_i\left[1 - P^*\right].$$

Taking the ratios of v_{1i} and v_{3i} we get

$$\frac{v_{1i}}{v_{3i}} = \begin{cases} \frac{1-P^*P_A}{1-P^*} & C_i = AB \\ \frac{1-P^*P_A}{1-P^*} + \frac{\lambda_{B\cdot}^e}{\delta_i[1-P^*]} & C_i = BA. \end{cases}$$

Given that we have shown that the joint distribution of v_i is identified then the distribution of v_{1i}/v_{3i} also must be identfified. If $P(AB) > 0$ then this distribution will have point mass at $\frac{1-P^*P_A}{1-P^*}$ that occurs with probability $P\left(AB|AB, BA\right)$. Note as well that since the support of δ_i is the real line, then v_{1i}/v_{3i} is strictly greater than $\left(1 - P^* P_A\right) / \left(1 - P^*\right)$ for the $C_i = BA$ types. In that case, $P(AB|AB, BA)$ and P^* are identified from the probability and the value at the minimum of the support (since we showed above that P_A is identified).[7] If instead $P(AB) = 0$ then we can use the ratios of v_{2i} and v_{3i} to identify P^* in a similar fashion.[8]

[7] As a practical matter in the estimation we use additional information as we observe the fraction of job-to-job transitions that are voluntary directly from survey data. Here we show that we can identify P^* without that knowledge.

[8] The only potential complication is the case in which either $P(AB) = 0$ or $P(BA) = 0$ and δ_i takes on only a single value. In that case both v_{1i}/v_{3i} and v_{2i}/v_{3i} will take only a single value, so from this alone we can not tell whether $P(AB) = 0$ or $P(BA) = 0$. However, we show that when we take into account the values of the identified hazard rates, we can tell which case we are in. To see this, suppose that were not the case and that the true model has $P\left(BA\right) = 0$ and let δ be the single value of δ_i. We would have to have another model with $P(AB) = 0$ and an

If both $P(BA) > 0$ and $P(AB) > 0$, since P^* is identified, we can identify the distribution of δ_i conditional on AB and BA as

$$\begin{aligned}Pr\left(\delta_i \leq d \mid C_i = AB\right) =& Pr\left(\frac{v_{3i}}{1-P^*} \leq d \mid \frac{v_{1i}}{v_{3i}} = \frac{1-P^*P_A}{1-P^*}\right)\\ Pr\left(\delta_i \leq d \mid C_i = BA\right) =& Pr\left(\frac{v_{3i}}{1-P^*} \leq d \mid \frac{v_{2i}}{v_{3i}} = \frac{1-P^*P_B}{1-P^*}\right).\end{aligned}$$

Let $Med(\cdot \mid \cdot)$ denote the conditional median (though any quantile will work) we know that

$$\begin{aligned}Med\left(v_{2i} \mid \frac{v_{1i}}{v_{3i}} = \frac{1-P^*P_A}{1-P^*}\right) =& Med\left(v_{2i} \mid C_i = AB\right)\\ =& Med\left(\delta_i\left[1-P^*P_B\right] + \lambda_A^e \mid C_i = AB\right)\end{aligned}$$

so we can write

$$\lambda_A^e = Med\left(v_{2i} \mid \frac{v_{1i}}{v_{3i}} = \frac{1-P^*P_A}{1-P^*}\right) - \left[1-P^*P_B\right] Med\left(\delta_i \mid \frac{v_{1i}}{v_{3i}} = \frac{1-P^*P_A}{1-P^*}\right).$$

we have shown that everything on the right hand side is identified, so λ_A^e must be as well.

From an analogous argument we can show

$$\lambda_B^e = Med\left(v_{1i} \mid \frac{v_{1i}}{v_{3i}} > \frac{1-P^*P_A}{1-P^*}\right) - \left[1-P^*P_A\right] Med\left(\delta_i \mid \frac{v_{1i}}{v_{3i}} > \frac{1-P^*P_A}{1-P^*}\right),$$

so λ_B^e is identified as well.

When $P(BA) = 0$, we can use the same approach to get $Pr\left(\delta_i \leq d \mid C_i = AB\right)$ and λ_A^e, but λ_B^e is not identified in this case. Likewise, when $P(AB) = 0$, we can use this approach to identify $Pr\left(\delta_i \leq d \mid C_i = BA\right)$ and λ_B^e, but λ_A^e is not identified. This is quite natural, since in the former case no one prefers B to A, so there is no way of identifying λ_B^e. Notice, that it is identified from wage data, since λ_B^e affects wages for both $C_i = BA$ and $C_i = B0$ type of

alternative value of $\widetilde{\delta}$, $\widetilde{P^*}$, and $\widetilde{\lambda_B^e}$ that satisfy the three equations

$$\begin{aligned}\widetilde{\delta}\left[1-\widetilde{P^*}P_A\right] + \widetilde{\lambda_B^e} =& \delta\left[1-P^*P_A\right]\\ \widetilde{\delta}\left[1-\widetilde{P^*}P_B\right] =& \delta\left[1-P^*P_B\right] + \lambda_A^e\\ \widetilde{\delta}\left[1-\widetilde{P^*}\right] =& \delta\left[1-P^*\right].\end{aligned}$$

But since $P_A + P_B = 1$, if we subtract the third equation from the first two we can show

$$\begin{aligned}\widetilde{\lambda_B^e} =& \left[\delta P^* - \widetilde{\delta}\widetilde{P^*}\right] P_B\\ \lambda_A^e =& \left[\widetilde{\delta}\widetilde{P^*} - \delta P^*\right] P_A,\end{aligned}$$

but then $\lambda_A^e > 0$ implies $\delta P^* < \widetilde{\delta}\widetilde{P^*}$ which implies that $\widetilde{\lambda_B^e} < 0$ which is not in the parameter space and thus we have a contradiction. Thus we can distinguish between these two cases and determine whether $P(AB) = 0$ or $P(BA) = 0$.

workers. Also, when we increase the number of job types, J, from two to more types, all that we require for identification of λ_k^e is that there are some individuals preferring k to other jobs that they would also take. This seems like a very reasonable assumption.

Identification of $P(AB), P(BA), P(0)$**,** $P(A0)$**, and** $P(B0)$

This is trivial given the previous result and infinite time. $P(0)$ is identified directly from the data as those that never work, $P(A0)$ as those that only work at an A type firm and $P(B0)$ as those that only work for a B type firm. Since we know $P(AB|AB, BA)$ and the probability of working both jobs, $P(AB)$ and $P(BA)$ are also identified.

Identification of the distribution of δ_i **conditional on** $C_i = A0$ **and on** $C_i = B0$

This is simpler than the cases above, since we can just take the survivor function of just a single spell for each of these. As long as $P(A0) > 0$ we can identify the duration for the first A firm type spell we observe (for everyone who would take an A type job). Let this value be T_i.

This is

$$\begin{aligned} Pr(T_i \geq t) = & Pr(T_i \geq t \mid A0)\, Pr(A0) + Pr(T_i \geq t \mid AB)\, Pr(AB) \\ & + Pr(T_i \geq t \mid BA)\, Pr(BA). \end{aligned}$$

We have shown everything in this expression other than $Pr(T_i \geq t \mid A0)$ is identified, so this term must also be identified. It is the conditional Laplace transform for the hazard rate out of the job for this group which is $\delta_i [1 - P^* P_A]$. Since $P^*$and P_A are identified, $Pr(\delta_i \leq d \mid A0)$ is as well.

The analogous argument gives $Pr(\delta_i \leq d \mid B0)$.

D.6 Theorem D.2

This proof is virtually identical the base model but we leave it essentially self contained. The main difference is that we include δ_i as part of the joint distribution we identify.

We need the following alternative assumption

Assumption D.2' *The econometrician observes*

1. *The full history of job type spells with start and stop dates as well as the value of j at each job.*
2. *If the individual is working, wages observed at the integers 1.0...,2.0..., for at least 8 periods.*

Theorem D.2 *Under Assumptions 1,D.2′, and 4-6 with the data generated by the model exposited in the general model section above, we can identify*

1. *The distribution of measurement error* ξ_{it}
2. *LBD human capital* ψ_1
3. *The joint distribution of* $(R_{iA00}, R_{iAB0}, \pi_{iA}, R_{iB00}, \pi_{iB}, R_{iA01}, R_{iAB1}, R_{iB01}, \delta_i)$ *conditional on* $C_i = AB$ *if* $Pr(C_i = AB) > 0$.
4. *The joint distribution of* $(R_{iA00}, \pi_{iA}, R_{iBA0}, R_{iB00}, \pi_{iB}, R_{iA01}, R_{iBA1}, R_{iB01}, \delta_i)$ *conditional on* $C_i = BA$ *if* $Pr(C_i = BA) > 0$.
5. *The joint distribution of* $(R_{iA00}, \pi_{iA}, R_{iA01}, \delta_i)$ *conditional on* $C_i = A0$ *if* $Pr(C_i = A0) > 0$.
6. *The joint distribution of* $(R_{iB00}, \pi_{iB}, R_{iB01}, \delta_i)$ *conditional on* $C_i = B0$ *if* $Pr(C_i = B0) > 0$.

Proof

To shorten some of the expressions we will use shorthand notation $\varpi_{ij\ell h_0 h}$ which we define as

$$\varpi_{ij\ell h_0 h} \equiv \log\left(R_{ij\ell h_0}\psi_h\right).$$

Identification of Distribution of Measurement Error (ξ_{it})

First, we identify the distribution of measurement error. We condition on a group who

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

We assume that the d_j's are sufficiently small (and non-negative), so spells do not overlap.

We can identify the joint distribution of (w_{i1}, w_{i2}) conditional on the events above for alternative values of d_1, d_2, and d_3.

Taking limits of the above object as $d_1 \downarrow 0, d_2 \downarrow 0$, and $d_3 \downarrow 0$, we can identify the conditional distribution of

$$\left(\varpi_{iA000} + \xi_{i1}, \varpi_{iA000} + \xi_{i2}\right),$$

for our conditioning group. Notice, that since $\psi_0 = 1$ then R_{iA00} is just the wage paid. Under assumption 5 using Kotlarski's lemma (Kotlarski 1967), we can identify the the marginal distributions of both the measurement error and ϖ_{iA000}.

Identification of λ_h

Next, we show that λ_h is identified. To economize on notation we will use $E(\cdot \mid d)$ to denote the expectation conditional on the events described above at values of $d = (d_1, d_2, d_3)$. We use the same conditioning group as in the Measurement Error section and continue to send $d_1 \downarrow 0$ and $d_3 \downarrow 0$, but allow d_2 to vary. This allows human capital to augment between period 1 and $1 + d_2$. We can identify the conditional characteristic function

$$\lim_{d_1,d_3 \downarrow 0} \frac{E\left(e^{\imath s w_{i2}} \mid d\right)}{\phi_{\xi}(s)} = \lim_{d_1,d_3 \downarrow 0} \left[e^{-\lambda_h d_2} E\left(e^{\imath s \varpi_{iA000}} \mid d\right) + \left(1 - e^{-\lambda_h d_2}\right) E\left(e^{\imath s \varpi_{iA011}} \mid d\right) \right].$$

By varying d_2 we can identify λ_h.[9] Intuitively, varying d_2 varies the time that the worker has to receive a human capital shock.

Identification of joint wage distribution for AB group

We now consider identification of the full wage distribution for the AB group conditional on δ_i. Identification is complicated, so to make this easier to follow we will do this in steps by showing identification of expanding subsets of the full distribution. We are implicitly assuming that $P(AB) > 0$ in what follows. If this is not the case we of course cannot identify the wage distribution for this group. One can use the same logic for the BA group exchanging A and B.

Conditioning set for Main Identification Result

For the AB types there are the seventeen different labor market statuses possible

[9] To see how, take the ratio of the derivatives of this function in terms of d_2 at two different values of d_2 and it will be a known function of λ_h. First, note that the derivative with respect to d_2 is

$$\lim_{d_1,d_3 \downarrow 0} \left[-\lambda_h e^{-\lambda_h d_2} E\left(e^{\imath s \varpi_{iA000}} \mid d\right) + \lambda_h e^{-\lambda_h d_2} E\left(e^{\imath s \varpi_{iA011}} \mid d\right) \right] = \lambda_h e^{-\lambda_h d_2} \left[E\left(e^{\imath s \varpi_{iA000}} \mid A\right) - E\left(e^{\imath s \varpi_{iA011}} \mid\mid A\right) \right],$$

where the notation $E(\cdot \mid A)$ means the expected value conditional on taking an A job first. Now take the ratio of this at two different values of d_2 say d_2^a and d_2^b then

$$\begin{aligned} \Delta(d_2^a, d_2^b) \equiv & \frac{\lambda_h e^{-\lambda_h d_2^a} \left[E\left(e^{\imath s \varpi_{iA000}} \mid A\right) - E\left(e^{\imath s \varpi_{iA011}} \mid\mid A\right) \right]}{\lambda_h e^{-\lambda_h d_2^b} \left[E\left(e^{\imath s \varpi_{iA000}} \mid A\right) - E\left(e^{\imath s \varpi_{iA011}} \mid A\right) \right]} \\ = & e^{\lambda_h \left(d_2^b - d_2^a\right)}. \end{aligned}$$

$\Delta(d_2^a, d_2^b)$ is directly identified from the data and

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

Table D3
Labor Market Statuses for AB workers

$j(i,t)$	$h(i,t)$	$\ell(i,t)$	$h_0(i,t)$	Wage	Log(Wage)
A	0	0	0	R_{iA00}	ω_{iA000}
A	0	B	0	R_{iAB0}	ω_{iAB00}
A	0	A	0	π_{iA}	ω_{iAA00}
A	1	0	0	$R_{iA00}\psi_1$	ω_{iA001}
A	1	B	0	$R_{iAB0}\psi_1$	ω_{iAB01}
A	1	A	0	$\pi_{iA}\psi_1$	ω_{iAA01}
A	1	0	1	$R_{iA01}\psi_1$	ω_{iA011}
A	1	B	1	$R_{iAB1}\psi_1$	ω_{iAB11}
A	1	A	1	$\pi_{iA}\psi_1$	ω_{iAA11}
B	0	0	0	R_{iB00}	ω_{iB000}
B	0	B	0	π_{iB}	ω_{iBB00}
B	1	0	0	$R_{iB00}\psi_1$	ω_{iB001}
B	1	B	0	$\pi_{iB}\psi_1$	ω_{iBB01}
B	1	0	1	$R_{iB01}\psi_1$	ω_{iB011}
B	1	B	1	$\pi_{iB}\psi_1$	ω_{iBB11}
0	0	NA	NA	NA	NA
0	1	NA	NA	NA	NA

where $j(i,t)$ is the current job type, $h(i,t)$ is the current human capital, $\ell(i,t)$ is the outside option when wages were negotiated, and $h_0(i,t)$ is the level of human capital when wages were negotiated.

From Table D3 one can see that for an AB worker's wage depend on the joint distribution of eight objects (in addition to ψ_1)

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

The model is overidentified so there are multiple ways to show identification. We focus on a particular set of transitions and show identification by taking limits. We emphasize that this is sufficient to show identification, we do not think it is necessary. We assume that workers start their labor market career in non-employed and receive their first job at $1-d_1$. The following table shows the transition path.

Transition	Time
Start at A	$1 - d_1$
Move to non-employment	$1 + d_2$
Start at B	$2 - d_3$
Move to non-employment	$2 + d_4$
Start at B	$3 - d_5 - d_6$
Move to A	$3 - d_6$
Move to non-employment	$3 + d_7$
Start at A	$4 - d_8$
Move to non-employment	$4 + d_9$
Start at B	$6 - d_{10}$
Move to non-employment	$6 + d_{11}$
Start at B	$8 - d_{12} - d_{13}$
Move to A	$8 - d_{13}$
Still Employed	8
Start at A from non-employment	After 8
Start another job from non-employment	After 8

with $d_j \geq 0$ for $j = 1, ..., 13$. We also assume that the d_j's are sufficiently small such that the above spells do not overlap. The goal here will be to look at the joint distribution of wages conditional on the d_j's. Analogous to above, we use the notation $E[\cdot \mid d]$ to mean the conditional expectation conditioning on events occurring at times denoted by $d_1 - d_{13}$.

Identification of Distribution of $(w_{i1}, ..., w_{i8})$ conditional on $(d, C_i = AB, \delta_i)$.

In going forward, we condition on wages from the first eight periods $(w_{i1}, ..., w_{i8})$. The last two spells will be analogous to the first and third type of spells we use in the first part of Theorem 1. Let Y_{1i} be the duration of the job spell at the A type firm for the first spell after period 8. This can end either in a transition to a B type firm or to non-employment. To mirror the notation in Theorem 1 let Y_{3i} the the duration of the last employment spell, i.e. from hiring until non-employment. From the last two spells in the transition table we use only the duration. Using well known results (see e.g. (French and Taber 2011)), we can write these durations as

$$\begin{aligned}
\log(Y_{1i}) &= v_{1i} + \omega_{1i} \\
\log(Y_{3i}) &= v_{3i} + \omega_{3i},
\end{aligned}$$

where v_{1i} and v_{3i} are the hazards from the two spells, and ω_{1i} and ω_{3i} have extreme value distribution. Let $\phi_\omega(t)$ be the characteristic function of the extreme value distribution, then we can identify the characteristic function of $(w_{i1}, ..., w_{i8}, v_{1i}, v_{3i})$ as

$$\frac{E\left[\exp\left(i\left(s_1w_{i1}+s_2w_{i2}+s_3w_{i3}+s_4w_{i4}+s_5w_{i5}+s_6w_{i6}+s_5w_{i5}+s_6w_{i6}+s_7w_{i7}+s_8w_{i8}+s_9Y_{1i}+s_{10}Y_{3i}\right.\right.\right.}{\phi_\omega\left(s_9\right)\phi_\omega\left(s_{10}\right)}.$$

$$\frac{\left.\left.\left.+s_6w_{i6}+s_7w_{i7}+s_8w_{i8}+s_9Y_{1i}+s_{10}Y_{3i}\right)\right)\mid d\right]}{\phi_\omega\left(s_9\right)\phi_\omega\left(s_{10}\right)}.$$

Since this characteristic function is identified then the joint distribution of $(w_{i1},...,w_{i8},v_{1i},v_{3i})$ is identified. As shown in the proof of Theorem 1, we know $C_i = AB$ when

$$\frac{v_{1i}}{v_{3i}} = \frac{1-P^*P_A}{1-P^*},$$

and all the terms on the right hand side are identified. We also know that in this case $v_{i1} = \delta_i\left[1-P^*P_A\right]$ and$[1-P^*P_A]$ is identified, so the joint distribution of $(w_{i1},...,w_{i8})$ conditional on d, δ_i and $C_i = AB$ is identified from the joint distribution of $(w_{i1},...,w_{i8},v_{1i},v_{3i})$ conditional on d as is the distribution of $(w_{i1},...,w_{i8})$ conditional on d, $v_{1i} = \delta_i\left[1-P^*P_A\right]$ and $v_{3i} = \delta_i\left[1-P^*\right]$.

While in principle we could show full identification of the eight dimensional distribution all at once, it is very complicated so instead we show it in pieces. We start with 3 parts.

Identification of joint distribution of $(R_{iA00}, R_{iB00}, R_{iAB0})$ for the AB types

We start by sending $d_1...d_6 \downarrow 0$ and look at the joint distribution of (w_{i1},w_{i2},w_{i3}). A complication is that at time $3-d_6$ individuals who moved directly from B to A could have either have gotten an outside offer from an A firm or been laid off and found a new job at an A firm immediately. Define $\rho_3(d)$ to be the probability that it is a voluntary transition. This a complicated but known expression since it involves only transition parameters, which we have shown are identified.

Then for any values of s_1-s_3 we can identify

$$\begin{aligned}
&\lim_{d_1...,d_6\downarrow 0}\frac{E\left[\exp\left(i\left(s_1w_{i1}+s_2w_{i2}+s_3w_{i3}\right)\right)\mid d,AB,\delta_i\right]}{\phi_\xi\left(s_1\right)\phi_\xi\left(s_2\right)\phi_\xi\left(s_3\right)}\\
=&\left[\lim_{d_1...,d_6\downarrow 0}\rho_3\left(d\right)\right]E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}\right)\right)\mid AB,\delta_i\right]\\
&+\left[\lim_{d_1...,d_6\downarrow 0}\left(1-\rho_3\left(d\right)\right)\right]E\left[\exp\left(i\left(\left(s_1+s_3\right)\varpi_{iA000}+s_2\varpi_{iB000}\right)\right)\mid AB,\delta_i\right]. \qquad \text{(E.2}'\text{.1)}
\end{aligned}$$

We will use the same basic argument for identification of the model throughout this section. We will be explicit about it here, but not as explicit in what follows (which will involve many more terms).

1. Letting $\lim_{d_1...,d_6\downarrow 0}$ identifies $\rho_3\left(d\right)$ as it is a known function of parameters that we have shown are identified.

2. By setting $s_3 = 0$ we can identify$E\left[\exp\left(i\left(s_1\varpi_{iA000} + s_2\varpi_{iB000}\right)\right) \mid AB\right]$ from the expression above.

3. Once this is identified, $E\left[\exp\left(i\left(s_1 + s_3\right)\varpi_{iA000} + s_2\varpi_{iB000}\right) \mid AB\right]$ is identified as we vary s_3.

4. Everything in the expression (E.2′.1) above is then identified except $E\left[\exp\left(i\left(s_1\varpi_{iA000} + s_2\varpi_{iB000} + s_3\varpi_{iAB00}\right)\right) \mid AB\right]$, so we can solve for this expression as well.

5. $E\left[\exp\left(i\left(s_1\varpi_{iA000} + s_2\varpi_{iB000} + s_3\varpi_{iAB00}\right)\right) \mid AB\right]$ is the characteristic function of $(\varpi_{iA000}, \varpi_{iB000}, \varpi_{iAB00})$, so since this is identified, the joint distribution of $(R_{iA00}, R_{iB00}, R_{iAB0})$ conditional on $C_i = AB$ and δ_i is identified.

Identification of $(R_{iA00}, R_{iB00}, R_{iAB0}, R_{A01}, R_{iB01}, R_{iAB1})$ for the AB types

Now, we extend the argument to include the joint distribution of

$$(R_{iA00}, R_{iB00}, R_{iAB0}, R_{A01}, R_{iB01}, R_{iAB1}),$$

for the AB types by adding wages from periods 4, 6, and 8. We will now vary d_7, which will allow for the possibility that human capital evolves between time 3 and $3 + d_7$ but send other values of d towards 0. There are 8 possible indistinguishable events that can occur in the data (after sending the other values of d arbitrarily close to zero); (the job-to-job transition to job A at time $3 - d_6$ is voluntary/involuntary) $\times$(human capital evolves or does not evolve between period 3 and $3 + d_7$)$\times$ (the job-to-job transition to job A at time $8 - d_{13}$ is voluntary/involuntary). Let ρ_3 and ρ_8 be the limit as $d_1..., d_6, d_8, ..., d_{13} \downarrow 0$ of the conditional probability that the job-to-job transitions are voluntary at time $3-d_6$ and $8 - d_{13}$, respectively. These are identified as they depend on transition parameters that we have shown are identified.

For any value of $s_1 - s_6$ we can identify

$$\begin{aligned}
&\lim_{d_1...,d_6,d_8,...,d_{13}\downarrow 0} \frac{E\left[\exp\left(i\left(s_1 w_{i1} + s_2 w_{i2} + s_3 w_{i3} + s_4 w_{i4} + s_5 w_{i6} + s_6 w_{i8}\right)\right) \mid d, AB, \delta_i\right]}{\phi_\xi(s_1)\phi_\xi(s_2)\phi_\xi(s_3)\phi_\xi(s_4)\phi_\xi(s_5)\phi_\xi(s_6)} \\
=&e^{-\lambda_h d_7}\left[\rho_3\rho_8\right] E\left[\exp\left(i\left(\left(s_1 + s_4\right)\varpi_{iA000} + \left(s_2 + s_5\right)\varpi_{iB000} + \left(s_3 + s_6\right)\varpi_{iAB00}\right)\right) \mid AB, \delta_i\right] \\
&+ e^{-\lambda_h d_7}\left[\rho_3\left(1 - \rho_8\right)\right] E\left[\exp\left(i\left(\left(s_1 + s_4 + s_6\right)\varpi_{iA000} + \left(s_2 + s_5\right)\varpi_{iB000} + s_3\varpi_{iAB00}\right)\right) \mid AB, \delta_i\right] \\
&+ e^{-\lambda_h d_7}\left[\left(1 - \rho_3\right)\rho_8\right] E\left[\exp\left(i\left(\left(s_1 + s_3 + s_4\right)\varpi_{iA000} + \left(s_2 + s_5\right)\varpi_{iB000} + s_6\varpi_{iAB00}\right)\right) \mid AB, \delta_i\right] \\
&+ e^{-\lambda_h d_7}\left[\left(1 - \rho_3\right)\left(1 - \rho_8\right)\right] E\left[\exp\left(i\left(\left(s_1 + s_3 + s_4 + s_6\right)\varpi_{iA000} + \left(s_2 + s_5\right)\varpi_{iB000}\right)\right) \mid AB, \delta_i\right] \\
&+ \left(1 - e^{-\lambda_h d_7}\right)\left[\rho_3\rho_8\right] E\left[\exp\left(i\left(s_1\varpi_{iA000} + s_2\varpi_{iB000} + s_3\varpi_{iAB00} + s_4\varpi_{iA011} + s_5\varpi_{iB011} + s_6\varpi_{iAB11}\right)\right) \mid AB, \delta_i\right] \\
&+ \left(1 - e^{-\lambda_h d_7}\right)\left[\rho_3\left(1 - \rho_8\right)\right] E\left[\exp\left(i\left(s_1\varpi_{iA000} + s_2\varpi_{iB000} + s_3\varpi_{iAB00} + \left(s_4 + s_6\right)\varpi_{iA011} + s_5\varpi_{iB011}\right)\right) \mid AB, \delta_i\right] \\
&+ \left(1 - e^{-\lambda_h d_7}\right)\left[\left(1 - \rho_3\right)\rho_8\right] E\left[\exp\left(i\left(\left(s_1 + s_3\right)\varpi_{iA000} + s_2\varpi_{iB000} + s_4\varpi_{iA011} + s_5\varpi_{iB011} + s_6\varpi_{iAB11}\right)\right) \mid AB, \delta_i\right] \\
&+ \left(1 - e^{-\lambda_h d_7}\right)\left[\left(1 - \rho_3\right)\left(1 - \rho_8\right)\right] E\left[\exp\left(i\left(\left(s_1 + s_3\right)\varpi_{iA000} + s_2\varpi_{iB000} + \left(s_4 + s_6\right)\varpi_{iA011} + s_5\varpi_{iB011}\right)\right) \mid AB, \delta_i\right].
\end{aligned}$$

We showed above that the first four expressions are identified. Thus we have four new expressions to identify:

$$
\begin{aligned}
&(a) E \exp\left[i\left(s_1 \varpi_{iA000} + s_2 \varpi_{iB000} + s_3 \varpi_{iAB00} + s_4 \varpi_{iA011} + s_5 \varpi_{iB011} + s_6 \varpi_{iAB11}\right) \mid AB, \delta_i\right] \\
&(b) E \exp\left[i\left(s_1 \varpi_{iA000} + s_2 \varpi_{iB000} + s_3 \varpi_{iAB00} + (s_4 + s_6)\, \varpi_{iA011} + s_5 \varpi_{iB011}\right) \mid AB, \delta_i\right] \\
&(c) E \exp\left[i\left((s_1 + s_3)\, \varpi_{iA000} + s_2 \varpi_{iB000} + s_4 \varpi_{iA011} + s_5 \varpi_{iB011} + s_6 \varpi_{iAB11}\right) \mid AB, \delta_i\right] \\
&(d) E \exp\left[i\left((s_1 + s_3)\, \varpi_{iA000} + s_2 \varpi_{iB000} + (s_4 + s_6)\, \varpi_{iA011} + s_5 \varpi_{iB011}\right) \mid AB, \delta_i\right].
\end{aligned}
$$

We use the same approach as above. If we evaluate at $s_3 = s_6 = 0$ these expressions are the same and thus $E\left[i\left(s_1 \varpi_{iA000} + s_2 \varpi_{iB000} + s_4 \varpi_{iA011} + s_5 \varpi_{iB011} \mid AB\right)\right]$ is identified. This identifies (d) for any values of $s_1 - s_6$.

Again using the same type of argument, given (d), keeping $s_3 = 0$ but varying the other values of s_j identifies (c) and setting $s_6 = 0$ but varying the others gives (b). Then everything in the large equation above is identified other than (a), so it is identified by varying all values of s_j.

Identification of ψ_1

Next, we consider identification of ψ_1 which we can do from $E\left[\log\left(W_{i7}\right) \mid d, AB\right]$ alone. In order to do this we condition on $1 < d_{11} < 2$ so that we observe w_{i7} and we will vary d_2, but send the rest of the $d_j \downarrow 0$. For this case there are three possibilities involving human capital: human capital has not evolved before period 7, human capital evolves between time 1 and time $1 + d_2$, and human capital evolves between periods 6 and 7. In addition, for each of these cases workers may or may not have gotten an outside offer from a B type firm between period 6 and 7.

$$
\begin{aligned}
&\lim_{d_1, d_3, \ldots, d_{13} \downarrow 0} E\left[\log\left(w_{i7}\right) \mid d, AB, \delta_i\right] = \\
&e^{-\lambda_h[1+d_2]}\left(e^{-\lambda_B^e} E\left[\varpi_{iB000} \mid AB, \delta_i\right] + \left(1 - e^{-\lambda_B^e}\right) E\left[\varpi_{iBB00} \mid AB, \delta_i\right]\right) \\
&\quad + \left(1 - e^{-\lambda_h d_2}\right)\left(e^{-\lambda_B^e} E\left[\varpi_{iB011} \mid AB, \delta_i\right] + \left(1 - e^{-\lambda_B^e}\right)\left(E\left[\varpi_{iBB11} \mid AB, \delta_i\right]\right)\right) \\
&\quad + \left(e^{-\lambda_h d_2} - e^{-\lambda_h[1+d_2]}\right)\left(e^{-\lambda_B^e} E\left[\varpi_{iB001} \mid AB, \delta_i\right] + \left(1 - e^{-\lambda_B^e}\right) E\left[\varpi_{iBB11} \mid AB, \delta_i\right]\right) + E\left(\xi_{i7}\right) \\
=& e^{-\lambda_h[1+d_2]}\left(e^{-\lambda_B^e} E\left[\varpi_{iB000} \mid AB, \delta_i\right] + \left(1 - e^{-\lambda_B^e}\right) E\left[\log\left(\pi_{iB}\right) \mid AB, \delta_i\right]\right) \\
&\quad + \left(1 - e^{-\lambda_h d_2}\right)\left(e^{-\lambda_B^e} E\left[\varpi_{iB011} \mid AB, \delta_i\right] + \left(1 - e^{-\lambda_B^e}\right)\left(E\left[\log\left(\pi_{iB}\right) + \log\left(\psi_1\right) \mid AB, \delta_i\right]\right)\right) \\
&\quad + \left(e^{-\lambda_h d_2} - e^{-\lambda_h[1+d_2]}\right)\left(e^{-\lambda_B^e} E\left[\varpi_{iB000} + \log\left(\psi_1\right) \mid AB, \delta_i\right] + \left(1 - e^{-\lambda_B^e}\right) E\left[\log\left(\pi_{iB}\right) + \log\left(\psi_1\right) \mid AB, \delta_i\right]\right) \\
&\quad + E\left(\xi_{i7}\right) \\
=& e^{-\lambda_h d_2} e^{-\lambda_B^e} E\left[\varpi_{iB000} \mid AB, \delta_i\right] + \left(1 - e^{-\lambda_h d_2}\right) e^{-\lambda_B^e} E\left[\varpi_{iB011} \mid AB, \delta_i\right] + E\left(\xi_{i7}\right) \\
&\quad + \left(1 - e^{-\lambda_A}\right) E\left[\log\left(\pi_{iB}\right) \mid AB, \delta_i\right] + \left[\left(1 - e^{-\lambda_A}\right) + \left(e^{-\lambda_h d_2} - e^{-\lambda_h[1+d_2]}\right)\right] \log\left(\psi_1\right).
\end{aligned}
$$

Everything is identified in this expression except $E\left[\log\left(\pi_{iB}\right) \mid AB, \delta_i\right]$ and $\log\left(\psi_1\right)$, so by varying d_2 they can be separately identified.

Identification of $(R_{iA00}, R_{iAB0}, R_{iB00}, R_{iA01}, R_{iAB1}, R_{iB01}, \pi_{iA}, \pi_{iB})$ *conditional on* AB

Now we assume that $1 < d_{11} < 2$ and $1 < d_9 < 2$ so that we observe wages at all times 1,...,8. By varying d_7, we can identify the expected value of $f(w_{i1}, ..., w_{i8})$ conditional on d, δ_i, and human capital arriving between time 3 and $3 + d_7$ (write this conditioning as $H_{i4} = 1$).

We will send the rest of the d_j's to zero (other than d_7, d_9, and d_{11}). Since we condition on human capital arriving between period 3 and $3 + d_7$, we know that the wage in the first period will be approximately R_{iA00}, the second period R_{iB00}, the fourth R_{iA01},and the sixth R_{iB01}. As before for the third and the eighth period the wage can take two values depending on whether the job-to-job transition was voluntary or not (R_{iA00} or R_{iAB0} in 3 and R_{iA01}or R_{iAB1} in 8). For period 5 the wage can take 3 values depending on outside offers: either R_{iA01} if no outside offers, R_{iAB1} if an offer from a B type only, or π_{iA} if an offer from an A type. Similarly in period 7 the wage can take 2 values depending on whether there was no outside offer (R_{iB01}) or an outside offer from a B firm (π_{iB}).[10] This gives a total of $2 \times 2 \times 3 \times 2 = 24$ different possibilities.

Analogous to above we define ρ_3 and ρ_8 be the limit as $d_1, ..., d_6, d_8, d_{10}, d_{12}, d_{13} \downarrow 0$ of the conditional probability that the job-to-job transitions are voluntary at time $3-d_6$ and $8 - d_{13}$, respectively.

Putting this together can identify the complicated expression with the relevant 24 terms.

[10] Since we are considering AB types they could not have gotten an offer from an A firm or they would have left

$$
\begin{aligned}
&\lim_{d_1,\ldots,d_6,d_8,d_{10},d_{12},d_{13}\downarrow 0} \frac{E\left[\exp\left(i\left(s_1 w_{i1}+s_2 w_{i2}+s_3 w_{i3}+s_4 w_{i4}+s_5 w_{i5}+s_6 w_{i6}+s_7 w_{i7}+s_8 w_{i8}\right)\right) \mid d, AB, \delta_i, H_{i4}=1\right]}{\phi_\xi(s_1)\,\phi_\xi(s_2)\,\phi_\xi(s_3)\,\phi_\xi(s_4)\,\phi_\xi(s_5)\,\phi_\xi(s_6)\,\phi_\xi(s_7)\,\phi_\xi(s_8)} \\
&= e^{-(\lambda_A^e+\lambda_B^e)} e^{-\lambda_B^e} [\rho_3\rho_8]\, E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+(s_4+s_5)\,\varpi_{iA011}+(s_6+s_7)\,\varpi_{iB011}+s_8\varpi_{iAB11}\right)\right) \mid AB,\delta_i\right] \\
&+ e^{-(\lambda_A^e+\lambda_B^e)} \left(1-e^{-\lambda_B^e}\right) [\rho_3\rho_8]\, E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+(s_4+s_5)\,\varpi_{iA011}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}+s_8\varpi_{iAB11}\right)\right) \mid AB,\delta_i\right] \\
&+ \left(1-e^{-\lambda_A^e}\right) e^{-\lambda_B^e} [\rho_3\rho_8]\, E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+s_4\varpi_{iA011}+s_5\varpi_{iAA11}+(s_6+s_7)\,\varpi_{iB011}+s_8\varpi_{iAB11}\right)\right) \mid AB,\delta_i\right] \\
&+ \left(1-e^{-\lambda_A^e}\right)\left(1-e^{-\lambda_B^e}\right) [\rho_3\rho_8]\, E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+s_4\varpi_{iA011}+s_5\varpi_{iAA11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}+s_8\varpi_{iAB11}\right)\right) \mid AB,\delta_i\right] \\
&+ e^{-\lambda_A^e}\left(1-e^{-\lambda_B^e}\right) e^{-\lambda_B^e} [\rho_3\rho_8]\, E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+s_4\varpi_{iA011}+(s_6+s_7)\,\varpi_{iB011}+(s_5+s_8)\,\varpi_{iAB11}\right)\right) \mid AB,\delta_i\right] \\
&+ e^{-\lambda_A^e}\left(1-e^{-\lambda_B^e}\right)\left(1-e^{-\lambda_B^e}\right) [\rho_3\rho_8]\, E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+s_4\varpi_{iA011}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}+(s_5+s_8)\,\varpi_{iAB11}\right)\right) \mid AB,\delta_i\right] \\
&+ e^{-(\lambda_A^e+\lambda_B^e)} e^{-\lambda_B^e} [\rho_3(1-\rho_8)]\, E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+(s_4+s_5+s_8)\,\varpi_{iA011}+(s_6+s_7)\,\varpi_{iB011}\right)\right) \mid AB,\delta_i\right] \\
&+ e^{-(\lambda_A^e+\lambda_B^e)} \left(1-e^{-\lambda_B^e}\right) [\rho_3(1-\rho_8)]\, E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+(s_4+s_5+s_8)\,\varpi_{iA011}+s_7\varpi_{iBB11}+s_6\varpi_{iB011}\right)\right) \mid AB,\delta_i\right] \\
&+ \left(1-e^{-\lambda_A^e}\right) e^{-\lambda_B^e} [\rho_3(1-\rho_8)]\, E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+(s_4+s_8)\,\varpi_{iA011}+s_5\varpi_{iAA11}+(s_6+s_7)\,\varpi_{iB011}\right)\right) \mid AB,\delta_i\right] \\
&+ \left(1-e^{-\lambda_A^e}\right)\left(1-e^{-\lambda_B^e}\right) [\rho_3(1-\rho_8)]\, E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+(s_4+s_8)\,\varpi_{iA011}+s_5\varpi_{iAA11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\right)\right) \mid AB,\delta_i\right] \\
&+ e^{-\lambda_A^e}\left(1-e^{-\lambda_B^e}\right) e^{-\lambda_B^e} [\rho_3(1-\rho_8)]\, E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+(s_4+s_8)\,\varpi_{iA011}+s_5\varpi_{iAB11}+(s_6+s_7)\,\varpi_{iB011}\right)\right) \mid AB,\delta_i\right] \\
&+ e^{-\lambda_A^e}\left(1-e^{-\lambda_B^e}\right)\left(1-e^{-\lambda_B^e}\right) [\rho_3(1-\rho_8)]\, E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+(s_4+s_8)\,\varpi_{iA011}+s_5\varpi_{iAB11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\right)\right) \mid AB,\delta_i\right] \\
&+ e^{-(\lambda_A^e+\lambda_B^e)} e^{-\lambda_B^e} [(1-\rho_3)\,\rho_8]\, E\left[\exp\left(i\left((s_1+s_3)\,\varpi_{iA000}+s_2\varpi_{iB000}+(s_4+s_5)\,\varpi_{iA011}+(s_6+s_7)\,\varpi_{iB011}+s_8\varpi_{iAB11}\right)\right) \mid AB,\delta_i\right] \\
&+ e^{-(\lambda_A^e+\lambda_B^e)} \left(1-e^{-\lambda_B^e}\right) [(1-\rho_3)\,\rho_8]\, E\left[\exp\left(i\left((s_1+s_3)\,\varpi_{iA000}+s_2\varpi_{iB000}+(s_4+s_5)\,\varpi_{iA011}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}+s_8\varpi_{iAB11}\right)\right) \mid AB,\delta_i\right] \\
&+ \left(1-e^{-\lambda_A^e}\right) e^{-\lambda_B^e} [(1-\rho_3)\,\rho_8]\, E\left[\exp\left(i\left((s_1+s_3)\,\varpi_{iA000}+s_2\varpi_{iB000}+s_4\varpi_{iA011}+s_5\varpi_{iAA11}+(s_6+s_7)\,\varpi_{iB011}+s_8\varpi_{iAB11}\right)\right) \mid AB,\delta_i\right] \\
&+ \left(1-e^{-\lambda_A^e}\right)\left(1-e^{-\lambda_B^e}\right) [(1-\rho_3)\,\rho_8]\, E\left[\exp\left(i\left((s_1+s_3)\,\varpi_{iA000}+s_2\varpi_{iB000}+s_4\varpi_{iA011}+s_5\varpi_{iAA11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}+s_8\varpi_{iAB11}\right)\right) \mid AB,\delta_i\right].
\end{aligned}
$$

$$
\begin{aligned}
&+ e^{-\lambda_A^e}\left(1-e^{-\lambda_B^e}\right)e^{-\lambda_B^e}\left[(1-\rho_3)\,\rho_8\right]E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+(s_4+s_5)\,\varpi_{iA011}+(s_6+s_7)\,\varpi_{iB011}+s_8\varpi_{iAB11}\right)\right)\mid AB,\delta_i\right]\\
&+ e^{-\lambda_A^e}\left(1-e^{-\lambda_B^e}\right)\left(1-e^{-\lambda_B^e}\right)\left[(1-\rho_3)\,\rho_8\right]E\left[\exp\left(i\left((s_1+s_3)\,\varpi_{iA000}+s_2\varpi_{iB000}+s_4\varpi_{iA011}+(s_5+s_8)\,\varpi_{iAB11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\right)\right)\mid AB,\delta_i\right]\\
&+ \left(1-e^{-\lambda_A^e}\right)e^{-\lambda_B^e}\left[(1-\rho_3)\,(1-\rho_8)\right]E\left[\exp\left(i\left((s_1+s_3)\,\varpi_{iA000}+s_2\varpi_{iB000}+(s_4+s_5+s_8)\,\varpi_{iA011}+(s_6+s_7)\,\varpi_{iB011}\right)\right)\mid AB,\delta_i\right]\\
&+ \left(1-e^{-\lambda_B}\right)e^{-(\lambda_A+\lambda_B)}\,(1-\rho_3)\,(1-\rho_8)\,E\left[\exp\left(i\left((s_1+s_3)\,\varpi_{iA000}+s_2\varpi_{iB000}+(s_4+s_5+s_8)\,\varpi_{iA011}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\right)\right)\mid AB,\delta_i\right]\\
&+ e^{-\lambda_B}\left(1-e^{-\lambda_A}\right)\left[(1-\rho_3)\,(1-\rho_8)\right]E\left[\exp\left(i\left((s_1+s_3)\,\varpi_{iA000}+s_2\varpi_{iB000}+(s_4+s_8)\,\varpi_{iA011}+s_5\varpi_{iAA11}+(s_6+s_7)\,\varpi_{iB011}\right)\right)\mid AB,\delta_i\right]\\
&+ \left(1-e^{-\lambda_A^e}\right)\left(1-e^{-\lambda_B^e}\right)\left[(1-\rho_3)\,(1-\rho_8)\right]E\left[\exp\left(i\left((s_1+s_3)\,\varpi_{iA000}+s_2\varpi_{iB000}+(s_4+s_8)\,\varpi_{iA011}+s_5\varpi_{iAA11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\right)\right)\mid AB,\delta_i\right]\\
&+ e^{-\lambda_A^e}\left(1-e^{-\lambda_B^e}\right)e^{-\lambda_B^e}\left[(1-\rho_3)\,(1-\rho_8)\right]E\left[\exp\left(i\left((s_1+s_3)\,\varpi_{iA000}+s_2\varpi_{iB000}+(s_4+s_8)\,\varpi_{iA011}+s_5\varpi_{iAB11}+(s_6+s_7)\,\varpi_{iB011}\right)\right)\mid AB,\delta_i\right]\\
&+ e^{-\lambda_A^e}\left(1-e^{-\lambda_B^e}\right)\left(1-e^{-\lambda_B^e}\right)\left[(1-\rho_3)\,(1-\rho_8)\right]E\left[\exp\left(i\left((s_1+s_3)\,\varpi_{iA000}+s_2\varpi_{iB000}+(s_4+s_8)\,\varpi_{iA011}+s_5\varpi_{iAB11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\right)\right)\mid AB,\delta_i\right].
\end{aligned}
$$

We now have sixteen new terms that have not been previously identified.

$$(a) E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+(s_4+s_5)\varpi_{iA011}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}+s_8\varpi_{iAB11}\right)\right)\mid AB,\delta_i\right]$$
$$(b) E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+s_4\varpi_{iA011}+s_5\varpi_{iAA11}+(s_6+s_7)\varpi_{iB011}+s_8\varpi_{iAB11}\right)\right)\mid AB,\delta_i\right]$$
$$(c) E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+s_4\varpi_{iA011}+s_5\varpi_{iAA11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}+s_8\varpi_{iAB11}\right)\right)\mid AB,\delta_i\right]$$
$$(d) E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+s_4\varpi_{iA011}+(s_6+s_7)\varpi_{iB011}+(s_5+s_8)\varpi_{iAB11}\right)\right)\mid AB,\delta_i\right]$$
$$(e) E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+(s_4+s_5+s_8)\varpi_{iA011}+(s_6+s_7)\varpi_{iB011}\right)\right)\mid AB,\delta_i\right]$$
$$(f) E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+(s_4+s_8)\varpi_{iA011}+s_5\varpi_{iAA11}+(s_6+s_7)\varpi_{iB011}\right)\right)\mid AB,\delta_i\right]$$
$$(g) E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+(s_4+s_8)\varpi_{iA011}+s_5\varpi_{iAA11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\right)\right)\mid AB,\delta_i\right]$$
$$(h) E\left[\exp\left(i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_3\varpi_{iAB00}+(s_4+s_8)\varpi_{iA011}+s_5\varpi_{iAB11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\varpi_{iB}\right)\right)\mid AB,\delta_i\right]$$
$$(i) E\left[\exp\left(i\left((s_1+s_3)\varpi_{iA000}+s_2\varpi_{iB000}+(s_4+s_5)\varpi_{iA011}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}+s_8\varpi_{iAB11}\right)\right)\mid AB,\delta_i\right]$$
$$(j) E\left[\exp\left(i\left((s_1+s_3)\varpi_{iA000}+s_2\varpi_{iB000}+s_4\varpi_{iA011}+s_5\varpi_{iAA11}+(s_6+s_7)\varpi_{iB011}+s_8\varpi_{iAB11}\right)\right)\mid AB,\delta_i\right]$$
$$(k) E\left[\exp\left(i\left((s_1+s_3)\varpi_{iA000}+s_2\varpi_{iB000}+s_4\varpi_{iA011}+s_5\varpi_{iAA11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}+s_8\varpi_{iAB11}\right)\right)\mid AB,\delta_i\right]$$
$$(l) E\left[\exp\left(i\left((s_1+s_3)\varpi_{iA000}+s_2\varpi_{iB000}+s_4\varpi_{iA011}+(s_5+s_8)\varpi_{iAB11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\right)\right)\mid AB,\delta_i\right]$$
$$(m) E\left[\exp\left(i\left((s_1+s_3)\varpi_{iA000}+s_2\varpi_{iB000}+(s_4+s_5+s_8)\varpi_{iA011}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\right)\right)\mid AB,\delta_i\right]$$
$$(n) E\left[\exp\left(i\left((s_1+s_3)\varpi_{iA000}+s_2\varpi_{iB000}+(s_4+s_8)\varpi_{iA011}+s_5\varpi_{iAA11}+(s_6+s_7)\varpi_{iB011}\right)\right)\mid AB,\delta_i\right]$$
$$(o) E\left[\exp\left(i\left((s_1+s_3)\varpi_{iA000}+s_2\varpi_{iB000}+(s_4+s_8)\varpi_{iA011}+s_5\varpi_{iAA11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\varpi_{iB}\right)\right)\mid AB,\delta_i\right]$$
$$(p) E\left[\exp\left(i\left((s_1+s_3)\varpi_{iA000}+s_2\varpi_{iB000}+(s_4+s_8)\varpi_{iA011}+s_5\varpi_{iAB11}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\right)\right)\mid AB,\delta_i\right].$$

We use the same basic approach as above. When we set various values of s_j to zero we can identify the components. To see how to identify all of these terms, setting $s_3 = s_8 = s_5 = 0$ all of the terms simplify to either

$$E\exp\left[i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_4\varpi_{iA011}+s_6\varpi_{iB011}+s_7\varpi_{iBB11}\right)\mid AB,\delta_i\right],$$

or

$$E\exp\left[i\left(s_1\varpi_{iA000}+s_2\varpi_{iB000}+s_4\varpi_{iA011}+(s_6+s_7)\varpi_{iB011}\right)\mid AB,\delta_i\right].$$

However, we have already shown identification of latter of these terms, which means the former is identified. Identification of this gives identification of term (m). Using a similar argument, setting $s_3 = s_7 = s_8 = 0$ we can identify term (n). Given these setting $s_3 = s_5 = 0$ we can show that (p),(i) and (l) are identified. Setting $s_3 = s_8 = 0$ we can identify (o), $s_3 = s_7 = 0$ gives (j), $s_5 = s_8 = 0$ gives (e), and $s_7 = s_8 = 0$ gives (f). Now with these setting $s_3 = 0$ gives (k), $s_8 = 0$ gives (g), $s_5 = 0$ gives (a), (d), and (h), and $s_7 = 0$ gives (b). This leaves only term (c) which is identified by varying all 8 terms given knowledge of all the other terms. This is the characteristic function for the joint distribution. Thus, we have shown that the joint distribution of wages for type AB workers can be non-parametrically identified, since the characteristic function uniquely determines the distribution.

Identification of the Distribution of Wages for the Other Types

Using a symmetric argument reversing A and B we can show that the distribution of

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

conditional on $C_i = BA$.

Next, consider the $A0$ types. We will use an argument similar to above, though it will be much simpler as there are fewer labor market statuses to worry about.

Table D4
Labor Market Statuses for $A0$ workers

$j(i,t)$	$h(i,t)$	$\ell(i,t)$	$h_0(i,t)$	Wage	$\log(Wage)$
A	0	0	0	R_{iA00}	ϖ_{iA000}
A	0	A	0	π_{iA}	ϖ_{iAA00}
A	1	0	0	$R_{iA00}\psi_1$	ϖ_{iA000}
A	1	A	0	$\pi_{iA}\psi_1$	ϖ_{iAA01}
A	1	0	1	$R_{iA01}\psi_1$	ϖ_{iA011}
A	1	A	1	$\pi_{iA}\psi_1$	ϖ_{iAA11}

From Table D4 one can see that for an $A0$ worker wages depend on the joint distribution of just three objects (in addition to ψ_1)

$$(R_{iA00}, \pi_{iA}, R_{iA01}).$$

Since there are three objects to identify we only need to use the first three periods. We consider the following the transition path. People begin non-employed at time zero and we will take $d_4 > 1$

Transition	Time
Start at A	$1 - d_1$
Move to non-employment	$1 + d_2$
Start at A	$2 - d_3$
Move to non-employment	$2 + d_4$
Still Working	3
Spell starting at A type firm	after 3

Let T_i be the duration of the spell with the hazard rate v_i.
We can identify

$$\begin{aligned}\frac{E\exp\left[i\left(s_1w_{i1}+s_2w_{i2}+s_3w_{i3}+s_4T_i\right) \mid d\right]}{\psi_w\left(s_4\right)} =& P(AB \mid d)E\exp\left[i\left(s_1w_{i1}+s_2w_{i2}+s_3w_{i3}+s_4v_i\right) \mid AB,d\right]\\ &+P(BA \mid d)E\exp\left[i\left(s_1w_{i1}+s_2w_{i2}+s_3w_{i3}+s_3v_i\right) \mid BA,d\right]\\ &+P(A0 \mid d)E\exp\left[i\left(s_1w_{i1}+s_2w_{i2}+s_3w_{i3}+s_4v_i\right) \mid A0,d\right].\end{aligned}$$

Since everything else in this expression is identified, we can identify $E\exp\left[i\left(s_1w_{i1}+s_2w_{i2}+s_3w_{i3}+s_4v_i\right) \mid A0,d\right]$.

Furthermore, analogous to the argument above using d_7, we now vary d_2 to identify the expected value of $(w_{i1}, w_{i2}, w_{i3}, \delta_i)$ conditional on d and human capital arriving between time 1 and $1 + d_2$ (write this conditioning as $H_{i2} = 1$). Then we can identify

$$E\exp\left[i\left(s_1w_{i1}+s_2w_{i2}+s_3w_{i3}+s_4\delta_i\left[1-P^*P_A\right]\right) \mid A0,d,H_{i2}=1\right].$$

In this case there is only one thing to worry about - whether the worker received and offer from another A firm between periods 2 and 3. Thus taking $d_1 \downarrow 0$ and $d_3 \downarrow 0$ we can identify

$$\begin{aligned}
&\lim_{d_1,d_3\downarrow 0} \frac{E\exp\left[i\left(s_1 w_{i1}+s_2 w_{i2}+s_3 w_{i3}\right) \mid A0, d, \delta_i, H_{i2}=1\right]}{\phi_\xi\left(s_1\right)\phi_\xi\left(s_2\right)\phi_\xi\left(s_3\right)} \\
=&e^{-\lambda_A} E\exp\left[i\left(s_1\varpi_{iA000}+\left[s_2+s_3\right]\varpi_{iA011}\right) \mid A0, \delta_i\right] \\
&+\left(1-e^{-\lambda_A}\right) E\exp\left[i\left(s_1\varpi_{iA000}+s_2\varpi_{iA011}+s_3\varpi_{iAA11}\right) \mid A0, \delta_i\right].
\end{aligned}$$

Set $s_3 = 0$ and we can identify $E\exp\left[i\left(s_1\varpi_{iA000}+s_2\varpi_{iA011}\right) \mid A0, \delta_i\right]$. Knowledge of this gives knowledge of $E\exp\left[i\left(s_1\varpi_{iA000}+\left[s_2+s_3\right]\varpi_{iA011}\right) \mid A0, \delta_i\right]$ and then allowing s_3 to vary means we can identify $E\exp\left[i\left(\varpi_{iA000}+s_2\varpi_{iA011}+s_3\varpi_{iAA11}\right) \mid A0, \delta_i\right]$ and thus the joint distribution of $\left(R_{iA00}, \pi_{iA}, R_{iA01}\right)$ conditional on δ_i for $C_i = A0$.

An analogous argument gives identification of the joint distribution of $\left(R_{iB00}, \pi_{iB}, R_{iB01}\right)$ conditional on δ_i for $C_i = B0$.

Thus, we have shown that wages, turnover parameters, and type proportions are identified.

Appendix E Robustness and Identification in Practice

Here we provide more details on the exercises discussed in Section 9 of the paper. In Section 9, we discussed five different checks. The main results were presented in the paper, but the details of four of the checks were not. We present the details in this section. We also present some more details on the exercise in Section 9.1.

E.1 Alternative Auxiliary Parameters

We discuss this in more detail in the main text of the paper in Section 9.1, but here we show how the estimated parameter values differ across the specifications.

Table E.1: Estimation under Alternative Auxiliary Parameters: Sensitivity of Structural Parameters

		Baseline	Alternative 1	Alternative 2	Alternative 3
E_θ	Mean worker productivity	4.263	3.1%	0.9%	3.6%
σ_θ	Std dev of worker productivity	0.217	1.6%	1.0%	-1.9%
σ_{v^p}	Std dev of match productivity	0.212	-5.7%	-0.6%	-0.8%
α	Weight on log wage	3.574	-39.5%	88.7%	2.7%
β	Bargaining power	0.848	2.2%	2.1%	-2.6%
P^*	Probability of imm offer upon job dest	0.391	-4.2%	6.0%	-1.3%
λ^n	Non-employment job offer arrival rate	0.993	-22.6%	0.4%	-7.2%
λ^e	Employment job offer arrival rate	2.076	-36.2%	-4.6%	-3.7%
μ_δ	Mean of log job destruction dist	-2.950	-17.1%	-1.2%	0.6%
σ_δ	Std dev of log job destruction dist	2.273	-31.0%	-6.6%	-0.1%
$b_1 \times 100$	Coef on linear term (human capital)	0.209	-468.1%	181.6%	36.1%
$b_2 \times 100$	Coef on quadratic term (human capital)	0.094	117.0%	-32.2%	26.3%
σ_ξ	Std dev of measurement error	0.139	-0.2%	0.7%	0.0%
f_u	Firm utility parameter	2.185	-25.9%	155.4%	-17.9%
f_p	Firm productivity parameter	0.141	2.1%	-11.9%	-5.3%
$f_{u,p} \times 100$	Firm utility $\times$ productivity parameter	0.370	1709.9 %	-6867.5%	-3746.1 %
σ_v	Std dev of idiosyncratic non-emp utility	0.352	52.7 %	11.1 %	-13.3 %
γ_θ	Worker ability cont to flow utility	-0.274	72.1 %	-6.8 %	-141.4 %

E.2 Estimation Under Alternative Restricted Models

In this subsection, we discuss our estimates of restricted versions of the main model. In each model we eliminate one of the proposed drivers of wage variance and see how this affects the fit of the model, the structural parameters obtained, and most importantly the counter-factual scenarios.[11]

[11] For search frictions we do not quite eliminate it, but make it much weaker.

Table E.2 show the results from seven different restrictions on the model. Given the computational time in doing this exercise, we use fewer worker simulations. In our base estimates we simulate 1,580,000 worker histories, but in these cases we use 158,000. The simulation error is small relative to the differences in the counterfactuals. For this reason, the results for the base model do not exactly correspond to the estimates in the main results in the paper, but one can see they are very close. In the rows with the auxiliary parameters, the first column shows the data values and then each column show the percentage difference between the real and simulated data starting with the baseline estimation. In the rows with the structural parameters, the column with the baseline model shows the parameters obtained previously and all other rows show the percentage difference. Finally, the counterfactuals are presented in the bottom of the table.

In the first restricted model in Table E.2, we eliminate human capital accumulation via learning by doing ($b_1 = b_2 = 0$). This model fits well except for the parameters of the estimated wage equation (experience, experience^2, and tenure^2). The counterfactuals are very similar. This is not surprising. Human capital operates fairly orthogonal to the other mechanisms in the model.[12] We next eliminate bargaining by setting $\beta = 1$. In this case, surprisingly, we can actually fit the tenure squared auxiliary parameter, but one can see that this is done by changing the human capital parameters (b_1 and b_2) such that we do not match those auxiliary parameters. This makes very little difference to the counterfactuals in the end. The fact that this fit is relatively good is in large part because the coefficient on tenure squared is relatively small. This means that we can fit it in a model without bargaining. This is also clear from the next subsection, where we measure the sensitivity of the auxiliary and counterfactuals to the structural parameters, the link between the bargaining parameter and the coefficient on tenure squared is clear.

Eliminating premarket skills across jobs ($\sigma_{v^p} = 0$) causes the model to miss wildly on the between job variance. The reason that the model does not generate more between job variance by increasing f_p is that this would cause $E(\widetilde{w}_{it}\widetilde{w}_{-it})$ to overshot. Eliminating all variation in premarket skills ($\gamma_\theta = \sigma_\theta = \sigma_{v^p} = 0$), the model misses in both the between job and between person variances.

Next, we eliminate preferences for non-pecuniary aspects of jobs. First, we set the variance of the prefence across workers within a job to zero by setting $var(v^u_{ij}) = 0$. In the baseline estimation α was estimated instead of $var(v^u_{ij})$, which we normalized to 1. We now fix $\alpha = 1$ instead. This leads the model to miss in many dimensions. The two most important ones

[12]Except that is of course interacts with frictions, since workers need to be employed to accumulate it.

Table E.2: Estimation under Alternative Models: Sensitivity of Auxiliary Parameters, Structural Parameters, and Counterfactuals

		Base Model	No LBD	No Monop.	No Pre. Across	No Pre. Total	No Tastes Across	No Tastes Total	$\lambda^n = 10.0$ $\lambda^e = 10.0$
Avg. Length Emp. Spell	377	0.1%	-2.5%	2.1%	-0.3%	-7.5%	11.6%	2.4%	10.4%
Avg. Length Nonnemp. Spell	91.4	-0.3%	-1.8%	1.6%	-0.6%	4.7%	-34.9%	-37.5%	-34.4%
Avg. Length Job	106	-0.0%	-1.3%	-1.2%	-2.4%	-10.9%	-7.2%	-16.7%	-26.0%
Sample mean $w_{i\ell jt}$	4.50	0.0%	0.0%	-0.1%	0.0%	-1.1%	2.1%	-0.4%	0.0%
Between Persons$\times$100	8.03	0.2%	-5.0%	0.5%	-0.5%	-81.5%	-18.6%	-23.3%	5.5%
Between Jobs$\times$100	2.87	-0.2%	-1.7%	-0.4%	-58.7%	-55.6%	-10.5%	-13.2%	-10.7%
Within Job$\times$100	1.49	0.1%	0.5%	-0.5%	9.7%	8.6%	7.3%	-0.1%	2.9%
$E\left(\widetilde{w}_{i\ell j}\widetilde{w}_{-i\ell j}\right) \times 100$	0.77	0.1%	-3.0%	-21.2%	6.9%	9.9%	-24.3%	-41.2%	0.8%
$E\left(\widetilde{r}_{-i\ell j}\widetilde{w}_{i\ell j}\right) \times 100$	0.69	-0.1%	3.7%	-5.5%	0.4%	14.1%	8.7%	33.3%	10.7%
$cov(\widetilde{r}_{-i\ell j}, \widetilde{S}_{i\ell j}) \times 100$	8.18	0.6%	-8.5%	0.6%	3.6%	-5.8%	-30.7%	-77.0%	-52.5%
Fraction Wage Drops	0.400	-1.8%	1.5%	4.2%	6.5%	4.2%	-27.1%	-41.2%	-9.8%
Coeff Exper$\times$100	2.48	-0.4%	-31.1%	-23.4%	6.2%	15.8%	9.7%	11.9%	17.2%
Coeff Exper$^2 \times 1000$	-0.291	-0.4%	-61.9%	16.2%	-0.2%	21.7%	-6.6%	-0.6%	6.0%
Coeff Tenure$^2 \times 1000$	-0.460	0.2%	34.0%	-2.2%	4.5%	-0.9%	28.5%	16.4%	6.1%
Var(Nonemployment)	16000	1.4%	6.7%	-1.2%	2.5%	1.7%	6.3%	17.3%	75.8%
Cov($\overline{w}_i$, Non-employment)	-3.42	-0.1%	5.1%	0.3%	-0.5%	-85.4%	16.2%	-1.0%	-5.3%
Var(Employment Dur)	102000	1.0%	0.4%	-0.3%	0.4%	-5.4%	2.7%	0.4%	2.7%
Invol Job to Job	0.205	0.0%	-0.2%	1.3%	0.5%	3.8%	-24.3%	-7.5%	7.7%
E_θ		4.263	1.9%	-2.9%	4.2%	2.2%	-2.3%	-6.6%	-6.6%
σ_θ		0.217	4.3%	4.0%	20.8%	-	-2.5%	-6.1%	7.5%
σ_{v^p}		0.212	2.0%	1.6%	-	-	-0.1%	0.0%	8.0%
α		3.574	-2.4%	-22.5%	-14.4%	-4.7%	-	-	35.0%
β		0.848	-7.6%	-	-3.1%	0.1%	-31.6%	-31.7%	-15.6%
P^*		0.391	0.4%	5.7%	2.2%	-0.2%	-0.0%	-0.0%	39.9%
λ^n		0.993	26.4%	-16.4%	-1.7%	-17.9%	146.8%	227.4%	-
λ^e		2.076	11.3%	1.2%	6.4%	5.2%	112.5%	134.8%	-
μ_δ		-2.950	-2.0%	-0.3%	-1.9%	-16.5%	20.1%	2.9%	16.5%
σ_δ		2.273	1.3%	-6.7%	-1.3%	-12.1%	-0.1%	-0.0%	3.8%
$b_1 \times 100$		0.209	-	2969.5%	-393.7%	-175.1%	19.0%	20.3%	27.7%
$b_2 \times 100$		0.094	-	-618.1%	92.1%	84.4%	-18.6%	9.6%	33.2%
σ_ξ		0.139	0.6%	2.6%	3.9%	6.2%	0.0%	0.0%	-0.2%
f_u		2.185	-6.1%	-11.8%	-8.2%	-20.9%	-90.7%	-	-16.3%
f_p		0.141	8.1%	-14.9%	-1.9%	-1.0%	-0.0%	-0.0%	43.7%
$f_{u,p} \times 100$		0.370	2783.0%	116.3%	-2419.8%	-2823.7%	0.2%	0.2%	4780.3%
σ_ν		0.352	-4.0%	29.1%	12.8%	6.2%	-86.5%	-86.5%	-6.0%
γ_θ		-0.274	16.7%	56.2%	-174.3%	-408.5%	36.5%	36.5%	365.3%
Full Model		0.105	0.100	0.104	0.087	0.022	0.088	0.082	0.106
No Learning by Doing		0.096	0.100	0.098	0.081	0.013	0.081	0.074	0.096
No LBD/Monopsony		0.093	0.096	0.098	0.077	0.008	0.071	0.065	0.091
No LBD/Monop./Premarket Across		0.049	0.050	0.053	0.077	0.008	0.032	0.025	0.041
No LBD/Monop./Premarket Total		0.008	0.008	0.006	0.008	0.008	0.006	0.006	0.010
No LBD/Monop./Search		0.086	0.090	0.093	0.078	0.007	0.059	0.050	0.085
No LBD/Monop./Nonpecuniary		0.087	0.091	0.091	0.074	0.004	0.068	0.065	0.084
No LBD/Monop./Pre. Across/Search		0.049	0.049	0.053	0.077	0.008	0.031	0.024	0.039
No LBD/Monop./Pre. Total/Search		0.007	0.006	0.005	0.007	0.007	0.005	0.004	0.007
No LBD/Monop./Pre. Across/NonPec.		0.048	0.049	0.052	0.074	0.004	0.031	0.025	0.039
No LBD/Monop./Pre. Total/Nonpec.		0.006	0.007	0.005	0.003	0.004	0.006	0.006	0.009
No LBD/Monop./Search/Nonpec.		0.061	0.068	0.066	0.072	0.000	0.054	0.050	0.070
No LBD/Monop./Pre. Across/Search/Nonpec.		0.047	0.047	0.051	0.072	0.001	0.030	0.024	0.038

are the fraction of wage drops and the correlation across workers in tastes of jobs $E(\widetilde{S}_{i\ell j}\widetilde{r}_{-i\ell j})$. Trying to fit these, the model misses on other auxiliary parameters as well. Including the non-employment rate and the correlation of wages with coworkers.

We next eliminate preferences for non-pecuniary aspects of jobs completely by setting the variances of firm and match specific non-pecuniary terms to zero ($f_u = var(v_{ij}^u) = 0$).[13] Thus, workers now chose jobs simply by maximize wages. In addition to the two auxiliary parameters from before the correlation between the preference for the job and wages, $E(\widetilde{w}_{it}\widetilde{r}_{-it})$, is off as is $E(\widetilde{w}_{it}\widetilde{w}_{-it})$, since the parameters that determine it are related to the first.

We cannot completely eliminate search as in that case all individuals would start at their favorite job immediately and never leave making the model not very interesting and making it impossible to even measure many of our auxiliary parameters. Instead, we just increase the value of the arrival rates to 10. As predicted we are way off on-the-job length auxiliary parameters, but the counterfactuals change very little.

E.3 Sensitivity of Auxiliary Parameters and Counterfactuals to Alternative Structural Parameters

In the following three tables (E.3a-E.3c) we change the structural parameters to alternative values to see what happens to the counterfactuals and the auxiliary parameters.

In each of the tables the first row shows the value we change the structural parameters to. The first column show the main specification estimated in section 8 of the paper.

[13] Again, we fix $\alpha = 1$.

Table E.3a: Sensitivity of Auxiliary Parameters and Counterfactuals to Alternative Parameters: I

Counterfactual/Auxiliary Parameter	Base Model	$(b_1, b_2) = (0, 0)$	$(b_1, b_2) = (0.1, 0)$	$\beta = 1.0$	$\beta = 0.75$	$\sigma_\theta = 0$	$\sigma_\theta = 0.4$	$\sigma_{v^p} = 0$	$\sigma_{v^p} = 0.4$
Avg. Length Emp. Spell	377	374	382	377	376	377	377	377	377
Avg. Length Nonnemp. Spell	91.2	103.9	53.9	89.1	92.7	74.9	128.2	98.1	80.7
Avg. Length Job	106	108	99	106	106	104	108	106	105
Sample mean $w_{i\ell jt}$	4.50	4.41	5.32	4.55	4.47	4.49	4.55	4.37	4.75
Between Persons×100	8.00	7.20	26.31	7.86	8.29	3.81	16.58	5.50	13.70
Between Jobs×100	2.88	2.59	8.66	2.66	3.29	2.86	2.94	1.02	7.01
Within Job×100	1.49	1.44	3.44	1.38	1.68	1.48	1.52	1.48	1.52
$E\left(\widetilde{w}_{i\ell j}\widetilde{w}_{-i\ell j}\right) \times 100$	0.77	0.73	1.76	0.76	0.78	0.76	0.79	0.85	0.74
$E\left(\widetilde{r}_{-i\ell j}\widetilde{w}_{i\ell j}\right) \times 100$	0.69	0.41	0.31	0.55	0.82	0.62	0.83	1.26	-0.15
$cov(\widetilde{r}_{-i\ell j}, \widetilde{S}_{i\ell j}) \times 100$	8.21	8.18	7.92	8.16	8.23	8.40	7.96	9.65	6.04
Fraction Wage Drops	0.392	0.415	0.277	0.411	0.384	0.394	0.389	0.413	0.360
Coeff Exper×100	2.47	1.22	12.57	1.17	3.42	2.48	2.47	2.35	2.71
Coeff Exper$^2 \times 1000$	-0.292	-0.083	-1.997	-0.208	-0.355	-0.297	-0.285	-0.283	-0.312
Coeff Tenure$^2 \times 1000$	-0.460	-0.446	-1.173	0.033	-0.820	-0.454	-0.481	-0.426	-0.532
Var(Nonemployment)	16150	20516	2924	15595	16543	7877	36539	20151	10411
Cov($\overline{w}_i$, Non-employment)	-3.43	-3.49	-1.93	-3.76	-3.14	0.03	-10.65	-4.13	-2.68
Var(Employment Dur)	102666	102535	102258	102634	102671	102771	102502	102584	102767
Invol Job to Job	0.205	0.199	0.237	0.208	0.203	0.203	0.210	0.206	0.203
Full Model	0.104	0.093	0.365	0.100	0.113	0.062	0.191	0.061	0.203
No Learning by Doing	0.096	0.093	0.127	0.093	0.104	0.054	0.184	0.055	0.190
No LBD/Monopsony	0.093	0.091	0.098	0.093	0.093	0.049	0.184	0.053	0.185
No LBD/Monop./Premarket Across	0.050	0.048	0.057	0.050	0.050	0.009	0.135	0.053	0.046
No LBD/Monop./Premarket Total	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008
No LBD/Monop./Search	0.086	0.085	0.086	0.086	0.086	0.037	0.193	0.054	0.138
No LBD/Monop./Non-pecuniary	0.087	0.086	0.091	0.087	0.087	0.041	0.183	0.050	0.168
No LBD/Monop./Pre. Across/Search	0.049	0.047	0.056	0.049	0.049	0.008	0.139	0.054	0.046
No LBD/Monop./Pre. Total/Search	0.007	0.006	0.007	0.007	0.007	0.007	0.007	0.006	0.007
No LBD/Monop./Pre. Across/NonPec.	0.048	0.047	0.055	0.048	0.048	0.007	0.134	0.050	0.046
No LBD/Monop./Pre. Total/Nonpec.	0.006	0.006	0.007	0.006	0.006	0.006	0.006	0.003	0.008
No LBD/Monop./Search/Nonpec.	0.061	0.064	0.058	0.061	0.061	0.010	0.180	0.048	0.082
No LBD/Mon./Pre./Search/Nonp. .	0.047	0.045	0.053	0.047	0.047	0.006	0.138	0.048	0.045

Table E.3b: Sensitivity of Auxiliary Parameters and Counterfactuals to Alternative Parameters: II

Counterfactual/Auxiliary Parameter	Base	$\alpha = 1.0$	$\alpha = 20$	$\lambda^e = 1.0$	$\lambda^e = 4.0$	$f_u = 0$	$f_u = 4$	$f_p = 0$	$f_p = 0.3$	$f_{u,p} = 0$	$f_{u,p} = 0.1$
Avg. Length Emp. Spell	377	377	380	365	382	378	376	377	377	377	377
Avg. Length Nonnemp. Spell	91.2	62.3	176.9	120.0	75.1	128.6	71.2	92.4	87.7	91.3	89.3
Avg. Length Job	106	102	107	137	86	107	104	106	106	106	106
Sample mean $w_{i\ell jt}$	4.50	4.28	4.72	4.48	4.53	4.58	4.44	4.48	4.57	4.50	4.51
Between Persons×100	8.00	9.50	7.19	7.81	8.22	7.62	8.51	7.57	9.14	8.00	8.00
Between Jobs×100	2.88	4.08	2.49	2.28	3.51	2.71	3.14	2.55	3.87	2.88	2.87
Within Job×100	1.49	2.08	1.43	1.53	1.50	1.46	1.56	1.49	1.50	1.49	1.50
$E\left(\widetilde{w}_{i\ell j}\widetilde{w}_{-i\ell j}\right) \times 100$	0.77	0.84	0.65	0.77	0.99	0.72	0.82	0.00	3.12	0.77	0.73
$E\left(\widetilde{r}_{-i\ell j}\widetilde{w}_{i\ell j}\right) \times 100$	0.69	0.09	1.15	-0.24	1.53	0.94	0.83	-0.03	2.68	0.67	1.09
$cov(\widetilde{r}_{-i\ell j}, \widetilde{S}_{i\ell j}) \times 100$	8.21	10.23	3.50	6.43	8.18	1.04	12.16	7.75	9.38	8.18	8.71
Fraction Wage Drops	0.392	0.430	0.278	0.404	0.388	0.372	0.394	0.396	0.382	0.393	0.386
Coeff Exper×100	2.47	5.28	1.86	1.99	2.46	2.09	3.07	2.44	2.54	2.46	2.50
Coeff Exper$^2 \times 1000$	-0.292	-0.504	-0.242	-0.264	-0.281	-0.260	-0.340	-0.290	-0.299	-0.291	-0.295
Coeff Tenure$^2 \times 1000$	-0.460	-1.453	-0.262	-0.214	-0.493	-0.333	-0.689	-0.452	-0.483	-0.459	-0.473
Var(Non-employment)	16150	2644	64763	21889	12121	37761	5570	16811	14231	16194	15111
Cov($\overline{w}_i$, Non-employment)	-3.43	-1.53	-1.36	-4.41	-2.34	-3.22	-2.77	-3.54	-3.15	-3.43	-3.33
Var(Employment Dur)	102666	102967	102431	100964	102807	102525	102757	102648	102691	102666	102678
Invol Job to Job	0.205	0.192	0.220	0.266	0.174	0.213	0.199	0.205	0.204	0.205	0.205
Full Model	0.104	0.137	0.092	0.097	0.113	0.099	0.113	0.097	0.126	0.104	0.104
No Learning by Doing	0.096	0.128	0.082	0.089	0.105	0.091	0.104	0.089	0.116	0.096	0.096
No LBD/Monopsony	0.093	0.100	0.078	0.090	0.093	0.088	0.096	0.086	0.112	0.093	0.092
No LBD/Monop./Premarket Across	0.050	0.056	0.044	0.047	0.053	0.046	0.053	0.042	0.070	0.050	0.049
No LBD/Monop./Premarket Total	0.008	0.008	0.007	0.008	0.008	0.007	0.008	0.000	0.029	0.008	0.007
No LBD/Monop./Search	0.086	0.099	0.058	0.086	0.086	0.081	0.087	0.081	0.093	0.086	0.085
No LBD/Monop./Non-pecuniary	0.087	0.091	0.077	0.091	0.083	0.082	0.090	0.081	0.102	0.087	0.086
No LBD/Monop./Pre. Across/Search	0.049	0.056	0.043	0.047	0.052	0.044	0.052	0.042	0.060	0.049	0.048
No LBD/Monop./Pre. Total/Search	0.007	0.008	0.004	0.007	0.007	0.006	0.007	0.000	0.017	0.007	0.006
No LBD/Monop./Pre. Across/Nonp.	0.048	0.055	0.044	0.047	0.051	0.045	0.052	0.042	0.061	0.048	0.048
No LBD/Monop./Pre. Total/Nonp.	0.006	0.007	0.006	0.007	0.006	0.006	0.006	0.000	0.019	0.006	0.006
No LBD/Monop./Search/Nonpec.	0.061	0.059	0.054	0.064	0.059	0.060	0.060	0.061	0.062	0.061	0.061
No LBD/Mon./Pre./Search/Nonp.	0.047	0.053	0.042	0.045	0.050	0.043	0.050	0.042	0.049	0.047	0.047

Table E.3c: Sensitivity of Auxiliary Parameters and Counterfactuals to Alternative Parameters: III

Counterfactual/Auxiliary Parameter	Base	$\lambda^n = 2.0$	$\sigma_\xi = 0$	$\mu_\delta = -\infty$	$\sigma_\delta = 0$	$E_\theta = 5.0$	$P^* = 0$	$\sigma_\nu = 0$	$\gamma_\theta = 0$
Avg. Length Emp. Spell	377	367	377	563	444	377	347	377	377
Avg. Length Nonnemp. Spell	91.2	68.7	91.2	82.6	82.4	91.2	91.6	66.8	84.8
Avg. Length Job	106	113	106	172	125	106	104	103	106
Sample mean $w_{i\ell jt}$	4.50	4.51	4.50	4.55	4.51	5.24	4.50	4.49	4.50
Between Persons×100	8.00	7.74	7.81	8.76	7.78	8.00	8.09	7.92	8.16
Between Jobs×100	2.88	2.56	2.48	1.60	2.83	2.88	2.71	2.87	2.87
Within Job×100	1.49	1.44	1.39	1.75	1.60	1.49	1.51	1.47	1.49
$E\left(\widetilde{w}_{i\ell j}\widetilde{w}_{-i\ell j}\right) \times 100$	0.77	0.74	0.77	0.81	0.84	0.77	0.76	0.75	0.76
$E\left(\widetilde{r}_{-i\ell j}\widetilde{w}_{i\ell j}\right) \times 100$	0.69	-0.03	0.69	2.06	1.49	0.69	0.90	0.58	0.66
$cov(\widetilde{r}_{-i\ell j}, \widetilde{S}_{i\ell j}) \times 100$	8.21	5.66	8.21	31.30	14.45	8.21	15.01	8.50	8.28
Fraction Wage Drops	0.392	0.404	0.381	0.356	0.390	0.392	0.359	0.395	0.393
Coeff Exper×100	2.47	2.08	2.45	2.12	2.42	2.47	2.44	2.50	2.47
Coeff Exper^2 × 1000	-0.292	-0.257	-0.290	-0.348	-0.269	-0.292	-0.302	-0.299	-0.293
Coeff Tenure^2 × 1000	-0.460	-0.295	-0.455	-0.132	-0.468	-0.460	-0.420	-0.459	-0.458
Var(Non-employment)	16150	11642	16150	16068	14791	16150	15730	4511	12799
Cov($\overline{w}_i$, Non-employment)	-3.43	-3.50	-3.43	-3.14	-3.48	-3.43	-3.58	-3.59	-2.76
Var(Employment Dur)	102666	100630	102666	108589	87956	102666	98665	102789	102705
Invol Job to Job	0.205	0.228	0.205	0.000	0.090	0.205	0.000	0.203	0.204
Full Model	0.104	0.098	0.104	0.102	0.103	0.104	0.104	0.103	0.106
No Learning by Doing	0.096	0.091	0.096	0.092	0.095	0.096	0.095	0.095	0.098
No LBD/Monopsony	0.093	0.091	0.093	0.091	0.092	0.093	0.092	0.093	0.094
No LBD/Monop./Premarket Across	0.050	0.047	0.050	0.050	0.050	0.050	0.050	0.049	0.051
No LBD/Monop./Premarket Total	0.008	0.008	0.008	0.007	0.008	0.008	0.008	0.008	0.008
No LBD/Monop./Search	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.086
No LBD/Monop./Non-pecuniary	0.087	0.090	0.087	0.074	0.082	0.087	0.086	0.086	0.087
No LBD/Monop./Pre. Across/Search	0.049	0.047	0.049	0.050	0.049	0.049	0.049	0.049	0.051
No LBD/Monop./Pre. Total/Search	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007
No LBD/Monop./Pre. Across/NonPec.	0.048	0.047	0.048	0.049	0.048	0.048	0.048	0.048	0.050
No LBD/Monop./Pre. Total/Nonpec.	0.006	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.006
No LBD/Monop./Search/Nonpec.	0.061	0.064	0.061	0.060	0.061	0.061	0.061	0.059	0.060
No LBD/Monop./Pre. /Search/Nonpec.	0.047	0.045	0.047	0.048	0.047	0.047	0.046	0.047	0.048

There are a lot of parameters, so we avoid a full discussion. Our view is that generally the identification map in Table 2 of the paper does quite well in giving an intuition of where identification is coming from. Some of the minor changes have to do with selection in terms of who works.

It should be pointed out that the identification map in Table 2 is overly simplified. This is because a lot of the parameters interact. For example, since human capital affects the relative taste for leisure when we eliminate human capital the moments related to duration of non-employment spells increase. This does not mean human capital helps identify this parameter in practice, though, because the arrival rate of jobs from non-employment would adjust to fit that moment.

E.4 Sensitivity of Results to Alternative "Normalization"

Recall, that in our main specification we made two "normalizations." However, given our functional form assumptions those are not precisely normalizations. Non-parametrically they would be, but once we have restricted the distribution of the error terms to be jointly normal, it is not formally innocuous. In this subsection, we show that if we use a very different "normalization" where we fix $\beta = 0$ and $\alpha = 1$ and then estimate $var(v_{ij}^u)$ together with $cov\left(v_{ij}^u, v_{ij}^p\right)$, we get the same main results, while the parameter estimates of course differ.

To understand the issues, consider our parametric model

$$\begin{aligned}
\log(\pi_{ij}) &= \theta_i + \mu_j^p + v_{ij}^p \\
u_{ij}(\pi_{ij}\psi_h) &= \alpha(\theta_i + \log(\psi_h)) + \left(\alpha\mu_j^p + \mu_j^u\right) + \left(\alpha v_{ij}^p + v_{ij}^u\right) \\
u_{i0t} &= \alpha\left[E(\theta_i) + \gamma_\theta\left(\theta_i - E(\theta_i)\right) + \nu_{io}^n\right].
\end{aligned}$$

If we ignored the bargaining (i.e. $\beta = 1$) part there are two normalizations that would be needed.

- A scale normalization on utility since utility is only identified up to scale. We essentially impose this in estimation by normalizing the variance of v_{ij}^u to be one.
- The second is to note that once we have the scale, we can identify

  $$cov\left(v_{ij}^p, \alpha v_{ij}^p + v_{ij}^u\right) = \alpha\sigma_p^2 + cov\left(v_{ij}^p, v_{ij}^u\right),$$

  so one can see that we cannot separately identify α from $cov\left(v_{ij}^p, v_{ij}^u\right)$. Loosely speaking, we have two pieces of information from the data. The first is wages from which we can identify π_{ij} and the second is revealed preferences from which we can identify the ranking of jobs. There are two reasons why the two rankings (productivity and utility) do not coincide. First, the strength of the covariance between the two in the joint offer distribution and second the value of α. We deal with this in the main specification by normalizing $cov\left(v_{ij}^{p}, v_{ij}^{u}\right) = 0$, so α is identified from this covariance (which has the nice interpretation as the relative weight one puts on wages relative to non-pecuniary aspects).

In the non-parametric framework, we showed that β is not formally identified. When we include bargaining some of these normalizations are not quite normalizations anymore, because we are assuming parametric distributions on the error terms and with bargaining wages are nonlinear. Intuitively, once we have made the "normalizations" above, β should be identified from the growth of wages on-the-job, since we have essentially pinned down the scale of the analogue of $V_{iA} - V_{i0}$. But intuitively, the value of β depends very much on these normalizations and a different normalization would lead to a different β.

To explore this, we try a different parameterization (which we think of as a normalization, but formally is not). We estimate the model fixing $\beta = 0$ and $\alpha = 1$, but we relax our other two assumptions by allowing the variance of v_{ij}^{u} to be free as well as $cov\left(v_{ij}^{p}, v_{ij}^{u}\right)$.

We use the specific parameterization

$$\begin{aligned}\log(\pi_{ij}) &= \theta_i + \mu_j^p + \sigma_p v_{ij}^p \\ u_{ij}(\pi_{ij}\psi_h) &= (\theta_i + \log(\psi_h)) + \left(\mu_j^p + \mu_j^u\right) + \left(a_1 v_{ij}^p + a_2 v_{ij}^u\right),\end{aligned}$$

where the $v's$ are both standard normal and uncorrelated with each other. Since they are jointly normal, it is a general way to do this. Thus, σ_p^2 is the variance of the productivity error term, $\sigma_p a_1$ is the covariance between the productivity and the taste error terms and $a_1^2 + a_2^2$ is the variance of the taste error term.

This proved to be a challenging exercise. In the base model, β was pinned down by the magnitude of the tenure effect (proxied by the coefficient on tenure squared). Setting $\beta = 0$ to get a similar low level of tenure effect requires a model in which workers are much closer to indifferent between all jobs. This requires the variance of the error term in the utility function $\left(a_1 v_{ij}^p + a_2 v_{ij}^u\right)$ to be very small. When that happens, the model is going to be sensitive to other parameters and hard to estimate. We simplified the model further by assuming $\gamma_\theta = 1$ and the variance of ν_{io}^{n} in the utility for non-employed to be zero and we no longer try to fit the variance of non-employment durations or the correlation between wages and non-employment durations. In practice, these parameters are very hard to pin down, but not important for the counterfactuals, so we do not lose much by ignoring them.

The second issue is human capital. This is another reason why one can get bargaining to be important. Since the value of human capital is higher on-the-job than off-the-job when high human capital workers get hired they get a very low initial wage leading to a substantial return to tenure when they subsequently get outside offers.

Our estimate of this model is presented in Table E.4.

We make a few comments about the results:

- The fit is not as good as in our base case. The main issue is human capital. The proposed model cannot fit both the human capital profile, which we understate, and the magnitude of the returns to tenure, which we overstate.
- The small values of a_1 and a_2 indicate the main part of the trade off. The variance of the v_{ij}^u terms is very low, so workers are close to being indifferent between the different jobs.
- Most importantly the main results of the counterfactual simulations are quite similar to our main model.
- The one odd exception is learning by doing. In this case, inequality goes up when we eliminate it. The reason is because we eliminate heterogeneity by allowing people to learn instantly, everyone has the maximum value. That increases inequality because it makes the bargaining more important (high human capital workers with an outside offer earn much more than high human capital workers directly hired from non-employment). If we eliminated human capital the other way-by assuming you never learn at all-we find that inequality decreases.

Table E.4: Sensitivity of Results to Alternative "Normalization"

	Data	Base Model	New Model
Avg. Length Emp. Spell	377	377	377
Avg. Length Nonnemp. Spell	91.4	91.2	95.9
Avg. Length Job	106	106	107
Sample mean $w_{i\ell jt}$	4.50	4.50	4.41
Between Persons$\times$100	8.03	8.00	7.60
Between Jobs$\times$100	2.87	2.88	3.04
Within Job$\times$100	1.49	1.49	1.42
$E\left(\widetilde{w}_{i\ell j}\widetilde{w}_{-i\ell j}\right) \times 100$	0.77	0.77	0.73
$E\left(\widetilde{r}_{-i\ell j}\widetilde{w}_{i\ell j}\right) \times 100$	0.69	0.69	0.69
$cov(\widetilde{r}_{-i\ell j}, \widetilde{S}_{i\ell j}) \times 100$	8.18	8.21	8.69
Fraction Wage Drops	0.400	0.392	0.391
Coeff Exper$\times$100	2.48	2.47	1.80
Coeff Exper$^2 \times 1000$	-0.291	-0.292	-0.190
Coeff Tenure$^2 \times 1000$	-0.460	-0.460	-0.602
Var(Employment Dur)	102000	102666	102518
Invol Job to Job	0.205	0.205	0.211
E_θ		4.263	4.259
σ_θ		0.217	0.203
σ_{v^p}		0.212	0.212
a_1			0.0059
a_2			0.0064
P^*		0.391	0.391
λ^n		0.993	0.567
λ^e		2.076	1.648
μ_δ		-2.950	-2.880
σ_δ		2.273	2.206
$b_1 \times 100$		0.209	0.424
$b_2 \times 100$		0.094	-0.021
σ_ξ		0.139	0.135
f_u		2.185	0.014
f_p		0.141	0.129
$f_{u,p}$ $\times$100		0.370	4.856
Full Model		0.105	0.102
No Learning by Doing		0.096	0.092
No LBD/Monopsony		0.093	0.092
No LBD/Monop./Premarket Across		0.049	0.049
No LBD/Monop./Premarket Total		0.008	0.007
No LBD/Monop./Search		0.086	0.080
No LBD/Monop./Nonpecuniary		0.087	0.085
No LBD/Monop./Pre. Across/Search		0.049	0.049
No LBD/Monop./Pre. Total/Search		0.007	0.007
No LBD/Monop./Pre. Across/NonPec.		0.048	0.049
No LBD/Monop./Pre. Total/Nonpec.		0.006	0.006
No LBD/Monop./Search/Nonpec.		0.061	0.051
No LBD/Monop./Pre. Across/Search/Nonpec.		0.047	0.046

E.5 Allowing for different degrees of complementarity between firms and workers

Our standard model of production is

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

Note that since this is logs, in levels it does impose some complementarity between firms and workers. However, since workers have log utility, this will have no impact on sorting. We do not explicitly estimate parameters on sorting, but instead perform a robustness check to see how the results would change in a model with more complementarity between worker skill $\left(\theta_i + v_{ij}^p\right)$ and the firm component μ_j^p. In particular, we generalize the production function to

$$\log\left(\pi_{ij}\right) = \overline{\theta} + \mu_j^p + \left(\omega \frac{\mu_j^p - \mu^\ell}{\mu^u - \mu^\ell} + (1-\omega)\right)\left(\theta_i - \overline{\theta} + v_{ij}\right)$$

where μ^u is the largest value of μ_j^p and $\mu^\ell$is the lowest. Note that when $\omega = 0$ we are in our base case, when $\omega = 1$ all workers are equally productive at the least productive firm.

We re-estimate the model with $\omega = 1/2$ and $\omega = 1$ and present the results in Table E.5. In the first panel, we show the fit of the model in terms of deviations from the targeted auxiliary parameters. One can see that both models fit very well. The next panel presents how the estimated parameters differ from the baseline model. In some cases the deviation is quite large-though the interpretation of some of the parameters are now very different.

Most important in the third panel we re-simulate the model decomposition. The main results of the model do not change very much. Premarket skills remain the most important-and the relative importance of the across variation in premarket skills is remarkably similar to the baseline case. The one thing that does change is the sign of removing search frictions, which go from a relatively small negative effect to a similar magnitude positive effect. This occurs almost by assumption-a positive α means that finding a good match is relatively more important for high θ_i workers than for low ones. This mean that relaxing search frictions helps those individuals. For a similar reason when $\alpha = 1$, eliminating preferences for non-pecuniary aspects of jobs also increases inequality. We view $\alpha = 1$ and $\alpha = 1/2$ as quite extreme parameter values, but ultimately this is an empirical question that we leave for future work.

Table E.5: Sensitivity of Results to Alternative Parameterization of Production Function

$$\log\left(\pi_{ij}\right) = \overline{\theta} + \mu_j + \left(\alpha \frac{\mu_j^p - \mu^\ell}{\mu^u - \mu^\ell} + (1-\alpha)\right)\left(\theta_i - \overline{\theta} + v_{ij}\right)$$

	Data	Baseline Model		
		$\alpha = 0$	$\alpha = \frac{1}{2}$	$\alpha = 1$
Avg. Length Emp. Spell	377	0.1%	0.3%	0.4%
Avg. Length Nonnemp. Spell	91.4	-0.3%	-1.9%	-1.0%
Avg. Length Job	106	-0.0%	-0.4%	-0.7%
Sample mean $w_{i\ell jt}$	4.50	0.0%	-1.5%	-3.4%
Between Persons$\times$100	8.03	0.2%	-0.0%	-3.2%
Between Jobs$\times$100	2.87	-0.2%	-2.2%	-3.1%
Within Job$\times$100	1.49	0.1%	-1.2%	0.1%
$E\left(\widetilde{w}_{i\ell j}\widetilde{w}_{-i\ell j}\right) \times 100$	0.77	0.1%	-2.7%	-0.3%
$E\left(\widetilde{r}_{-i\ell j}\widetilde{w}_{i\ell j}\right) \times 100$	0.69	-0.1%	0.7%	-1.8%
$cov(\widetilde{r}_{-i\ell j}, \widetilde{S}_{i\ell j}) \times 100$	8.18	0.6%	-0.5%	0.8%
Fraction Wage Drops	0.400	-1.8%	-1.8%	-1.6%
Coeff Exper$\times$100	2.48	-0.4%	-1.8%	-0.3%
Coeff Exper$^2 \times 1000$	-0.291	-0.4%	1.9%	0.4%
Coeff Tenure$^2 \times 1000$	-0.460	0.2%	-0.3%	-2.0%
Var(Nonemployment)	16000	1.4%	0.2%	2.0%
Cov($\overline{w}_i$, Non-employment)	-3.42	-0.1%	-0.4%	-1.1%
Var(Employment Dur)	102000	1.0%	1.0%	1.0%
Invol Job to Job	0.205	0.0%	0.2%	-0.9%
E_θ		4.263	-1.6%	-3.0%
σ_θ		0.217	18.7%	49.2%
σ_{v^p}		0.212	21.1%	35.0%
α		3.574	-0.0%	-0.0%
β		0.848	1.5%	0.0%
P^*		0.391	0.1%	0.3%
λ^n		0.993	-0.2%	-0.1%
λ^e		2.076	1.0%	4.4%
μ_δ		-2.950	0.0%	0.0%
σ_δ		2.273	-0.3%	-0.4%
$b_1 \times 100$		0.209	267.3%	-7.4%
$b_2 \times 100$		0.094	-52.2%	5.9%
σ_ξ		0.139	0.0%	0.0%
f_u		2.185	-1.8%	-0.0%
f_p		0.141	-42.0%	-80.4%
$f_{u,p}$ $\times$100		0.370	296.3%	170.3%
σ_v		0.352	-5.0%	-15.1%
γ_θ		-0.274	26.4%	89.2%
Full Model		0.105	0.104	0.101
No Learning by Doing		0.096	0.095	0.092
No LBD/Monopsony		0.093	0.093	0.089
No LBD/Monop./Premarket Across		0.049	0.047	0.041
No LBD/Monop./Premarket Total		0.008	0.003	0.000
No LBD/Monop./Search		0.086	0.100	0.115
No LBD/Monop./Nonpecuniary		0.087	0.092	0.098
No LBD/Monop./Pre. Across/Search		0.049	0.054	0.044
No LBD/Monop./Pre. Total/Search		0.007	0.002	0.000
No LBD/Monop./Pre. Across/NonPec.		0.048	0.057	0.044
No LBD/Monop./Pre. Total/Nonpec.		0.006	0.001	0.000
No LBD/Monop./Search/Nonpec.		0.061	0.075	0.102
No LBD/Monop./Pre. Across/Search/Nonpec.		0.047	0.062	0.044

References

French, E. and C. Taber (2011). Identification of models of the labor market. In D. C. O. Ashenfelter (Ed.), *The Handbook of Labor Economics Volume 4A*.

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