–NOT INTENDED FOR PUBLICATION–

Online Appendix for "Job Duration and Match Characteristics over the Business Cycle"

Ismail Baydur and Toshihiko Mukoyama

January 2020

A Estimation details

A.1 Coefficient estimates

Our dataset takes the following form. Let i denote an individual for $i = 1,..,N$, let j denote a particular job individual i holds for $j = 1,...,J_i$, and let t denote duration of the job for $t = 1,...,T_{ij}$. Because the unemployment rate changes every month, our dataset is represented in an extended form. In what follows, we denote vectors and matrices in bold to distinguish them from scalars. Accordingly, let $\mathbf{x}$ be the matrix representing our dataset. Then, each row, $\mathbf{x}_{ijt}$, corresponds to a monthly observation of individual i's j^{th} job at duration t.[33] Note that we also conveniently switch from the parentheses notation in the main text for t to a subscript notation to emphasize the discrete nature of the time-varying regressors. Also let k denote competing risks for $k = 1,...,K$, and let δ_{ijt}^k be an indicator for event type k of individual i's j^{th} job at duration t. Its value is zero for all competing risks except the last row. For the last row, at most one of δ_{ijt}^k is equal to 1. If the last row is also all 0's, the observation is right censored.

The partial log-likelihood for event k takes the following form:

$$\mathcal{L}_p^k \quad = \sum_{i=1}^{I}\sum_{j=1}^{J_i}\sum_{t=1}^{T_{ij}} \delta_{ijt}^k w_i \left\{ \mathbf{x}_{ijt}\boldsymbol{\beta^k} + o_i^k - \log\left[\sum_{m=1}^{I}\sum_{n=1}^{J_m} w_m \exp\left(\mathbf{x}_{mnt}\boldsymbol{\beta^k} + o_m^k\right)\right]\right\},$$

where w_i is the sampling weight and o_i^k is the cause-specific offset term, which is a variable with a coefficient equal to 1. The offset terms are useful for the frailty estimation. Otherwise, their values are set to 0. The cause-specific hazard model maximizes this log-likelihood separately for each competing event to get an estimate for the coefficients, $\hat{\boldsymbol{\beta}}^{\boldsymbol{k}}$.

The second term in the log-likelihood is the contribution of the risk set. Note that, for a given t, only the jobs spells that survived until t contribute to this term. For stratified

[33] More precisely, the observation for individual i's j^{th} job between duration $t-1$ and t, that is, $(t-1, t]$.

regressions, the summation for the contribution of the risk set runs over the job spells for that individual only; that is, it excludes the first summation sign. For subhazard regressions, we extend each job spell to the maximum duration observed in the dataset and multiply the sampling weights by the Kaplan-Meier estimate for censoring times. We then run a Cox regression to this extended dataset under either frailty or stratification specification.

The necessary condition for the maximization problem above satisfies the following score equation:

$$\frac{\partial \mathcal{L}_p^k}{\partial \hat{\boldsymbol{\beta}}^{\boldsymbol{k}}} \quad = \mathbf{J}(\hat{\boldsymbol{\beta}}^{\boldsymbol{k}}) = \mathbf{0}.$$

The covariance matrix for the coefficient estimates is equal to the inverse of the information matrix:

$$var(\hat{\boldsymbol{\beta}}^{\boldsymbol{k}}) \quad = \left[-\frac{\partial^2 \mathcal{L}_p^k}{\partial \hat{\boldsymbol{\beta}}^{\boldsymbol{k}} \partial \hat{\boldsymbol{\beta}}^{\boldsymbol{k}\prime}}\right]^{-1} = \left[\mathbf{F}(\hat{\boldsymbol{\beta}}^{\boldsymbol{k}})\right]^{-1}$$

For the stratified regressions, we report the standard errors clustered over individuals. They are calculated as follows:

$$var(\hat{\boldsymbol{\beta}}^{\boldsymbol{k}}) \quad = \left[\mathbf{F}(\hat{\boldsymbol{\beta}}^{\boldsymbol{k}})\right]^{-1} \left[\sum_{i=1}^{I} \mathbf{J_i}(\hat{\boldsymbol{\beta}}^{\boldsymbol{k}})' \mathbf{J_i}(\hat{\boldsymbol{\beta}}^{\boldsymbol{k}})\right] \left[\mathbf{F}(\hat{\boldsymbol{\beta}}^{\boldsymbol{k}})\right]^{-1},$$

where $\mathbf{J_i}(\hat{\boldsymbol{\beta}}^{\boldsymbol{k}})$ is the score vector summed over all the observations of individual i. For subhazard regressions, we correct the standard errors due to the Kaplan-Meier estimate by applying the procedure detailed in Fine and Grey (1999).

A.2 Post-estimation: Survival and cumulative incidence functions

The Cox model does not specify a functional form for the baseline hazard. After estimating $\boldsymbol{\beta}^k$, the baseline hazard function for cause k, $\hat{h}_0^k$, can be (non-parametrically) obtained as follows:

$$\hat{h}_0^k(t, \hat{\boldsymbol{\beta}}^{\boldsymbol{k}}) \quad = \frac{\sum_{i=1}^{I} \sum_{j=1}^{J_i} \delta_{ijt}^k w_i}{\sum_{i=1}^{I} \sum_{j=1}^{J_i} w_i \exp\left(\mathbf{x}_{ijt} \hat{\boldsymbol{\beta}}^{\boldsymbol{k}} + o_i^k\right)}.$$

To simplify the notation, we suppress the dependence of this function on $\mathbf{x}$ and δ^k. We calculate the survival function, S, and the cumulative incidence functions, P^k, for a hypothetical observation, $\mathbf{z_0}$. For this particular observation, define the following quantity:

$$\hat{h}(t, \hat{\boldsymbol{\beta}}^{\boldsymbol{k}} | \mathbf{z_0}) \quad = \sum_{k=1}^{K} \hat{h}_0^k(t, \hat{\boldsymbol{\beta}}^{\boldsymbol{k}}) \exp\left(\mathbf{z}_0 \hat{\boldsymbol{\beta}}^{\boldsymbol{k}}\right).$$

Then, the survival and the cumulative incidence functions for cause k are given by:

$$\begin{aligned}
\hat{S}(t, \hat{\boldsymbol{\beta}}^{\boldsymbol{k}} | \mathbf{z_0}) \quad &= \prod_{\tau=1}^{t} \left(1 - \hat{h}(\tau, \hat{\boldsymbol{\beta}}^{\boldsymbol{k}} | \mathbf{z_0})\right) \\
\hat{P}^k(t, \hat{\boldsymbol{\beta}}^{\boldsymbol{k}} | \mathbf{z_0}) \quad &= \sum_{\tau=1}^{t} \hat{S}(\tau, \hat{\boldsymbol{\beta}}^{\boldsymbol{k}} | \mathbf{z_0}) \hat{h}_0^k(\tau, \hat{\boldsymbol{\beta}}^{\boldsymbol{k}}) \exp\left(\mathbf{z}_0 \hat{\boldsymbol{\beta}}^{\boldsymbol{k}}\right).
\end{aligned}$$

For both of these functions, the parameter estimates non-linearly interact with each other. Because these estimates are random variables, the survival and cumulative incidence functions are also random. In the main text, we calculate the covariance matrix for the cumulative incidence functions by applying the method in Rosthøj, Anderson, and Abildstrom (2004).[34] Once the covariance matrix for $\hat{P}^k(.|\mathbf{z_0})$ is obtained, we use the Delta method to obtain the standard error for $\log(-\log(\hat{P}^k(.|\mathbf{z_0})))$.[35] This conversion guarantees that the confidence intervals are between 0 and 1. By back transformation, we calculate the confidence interval for $\hat{P}^k(.|\mathbf{z_0})$ as $[\exp(-\exp(U)), \exp(-\exp(L))]$, where L and U are the upper and lower bounds for $\log(-\log(\hat{P}^k(.|\mathbf{z_0})))$.

A.3 Frailty model and the EM algorithm

We used the survival package in R to estimate the cause-specific and subhazard models. This package implements a penalized likelihood approach. An alternative to the penalized likelihood approach is the Expectation-Maximization (EM) algorithm. We describe this method here to facilitate our discussion about estimating the non-parametric frailty model.[36]

Consider the normal frailty model where frailties for each competing event are coming from independent normal distributions with mean zero and variances θ^k to be estimated. The EM algorithm treats the frailty terms as unobserved data and estimates them via Bayes' rule. In implementation, the cause-specific hazard rates are estimated separately. The algorithm has two loops. The outer loop searches for the optimal variance by maximizing the *marginal*

[34]This method requires the covariance matrix for $\hat{\boldsymbol{\beta}}^{\boldsymbol{k}}$. Here, we included the covariance matrix for the frailty estimates as well. We used coxme package in R to obtain the standard errors for frailty terms.

[35]The calculated standard errors are equal to the ratio of the standard error of $\log(-\log(\hat{P}^k(.|\mathbf{z_0})))$ to $\hat{P}^k(.|\mathbf{z_0}) \left|\log(\hat{P}^k(.|\mathbf{z_0}))\right|$.

[36]See Therneau, Grambsch, and Pankratz (2003) for details.

full likelihood, for example, using the golden section search. The full likelihood contribution for individual i is given by:

$$
\begin{aligned}
L_i^f(\boldsymbol{\beta}, \mathbf{x}_i, \mathbf{o}_i) = \quad & \prod_{j=1}^{J_i} \prod_{t=1}^{T_{ij}} \prod_{k=1}^{K} \left[h_0^k(t, \boldsymbol{\beta^k}) \exp(\mathbf{x}_{ijt}\boldsymbol{\beta^k} + o_i^k) \right]^{\delta_{ijt}^k} \\
& \left[\exp(-h_0^k(t, \boldsymbol{\beta^k}) \exp(\mathbf{x}_{ijt}\boldsymbol{\beta^k} + o_i^k)) \right]^{1-\delta_{ijt}^k},
\end{aligned}
$$

where $\mathbf{o}_i$ is a vector of the frailty terms for individual i, and $\boldsymbol{\beta^k}$ is the vector of parameter estimates. More compactly, we can write this equation as:

$$
L_i^f(\boldsymbol{\beta}, \mathbf{x}_i, \mathbf{o}_i) = \prod_{k=1}^{K} L_i^{f,k}(\boldsymbol{\beta^k}, \mathbf{x}_i, o_i^k).
$$

We can define the marginal (full) likelihood contribution of individual i for event k by integrating out the offset terms for that event:

$$
L_i^{m,k}(\theta^k | \boldsymbol{\beta^k}, \mathbf{x}_i) = \int_{-\infty}^{\infty} L_i^{f,k}(\boldsymbol{\beta^k}, \mathbf{x}_i, o_i^k) f(o_i^k, \theta^k) do_i^k,
$$

where $f(o_i^k, \theta^k)$ is the normal density with variance θ^k. We stress two points here. First, the parameter vector and the frailty terms are all known at this stage because they are estimated in the inner loop we explain below. Therefore, the estimation is over θ^k only. Second, the marginal likelihood omits the likelihood contribution from other competing events. This treatment is valid under the independence assumption. More generally, we can write the likelihood contribution of individual i (after integrating out all the frailty terms) as the multiplication of the marginal full likelihood functions of all competing events:

$$
L_i^m(\boldsymbol{\theta} | \boldsymbol{\beta}, \mathbf{x}_i) = \prod_{k=1}^{K} L_i^{m,k}(\theta^k | \boldsymbol{\beta^k}, \mathbf{x}_i),
$$

where $\boldsymbol{\theta}$ is a vector of the variances for each of the frailty distributions. The maximization of the log of this objective function over all the individuals is equivalent to running cause-specific Cox regressions, because the parameters and the frailty terms do not interact with each other. When we relax the independence assumption in the non-parametric estimation, the integral is calculated over all the competing events and the objective function becomes:

$$L_i^m(\boldsymbol{\theta}|\boldsymbol{\beta}, \mathbf{x}_i) = \underbrace{\int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty}}_{\text{1 through K}} L_i^f(\boldsymbol{\beta^k}, \mathbf{x}_i, \mathbf{o}_i) f(\mathbf{o}_i, \boldsymbol{\theta}) do_i^K ... d_i^1.$$

In this case, $\boldsymbol{\theta}$ is a vector of parameters that governs the joint distribution of $\mathbf{o}_i$.

The inner loop estimates the Cox model for a given set of distributional parameters, and the frailty terms enter the estimation equation as offset terms. In the first iteration, we guess the frailty terms; for example, they are set equal to zero for all individuals. Once the Cox model is estimated, an update for frailty terms of individual i is obtained via Bayes' rule:

$$E(o_i^k|\mathbf{x}_i, \boldsymbol{\beta}^k, \theta^k) = \int_{-\infty}^{\infty} o_i^k \frac{L_i^{f,k}(\boldsymbol{\beta^k}, \mathbf{x}_i, o_i^k)}{L_i^{m,k}(\theta^k|\boldsymbol{\beta^k}, \mathbf{x}_i)} f(o_i^k, \theta^k) do_i^k.$$

Once again, this formulation assumes independence among the frailty terms. More generally, we can write this as follows:

$$E(o_i^k|\mathbf{x}_i, \boldsymbol{\beta}, \boldsymbol{\theta}) = \underbrace{\int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty}}_{\text{1 through K}} o_i^k \frac{L_i^f(\boldsymbol{\beta}, \mathbf{x}_i, \mathbf{o}_i)}{L_i^m(\boldsymbol{\theta}|\boldsymbol{\beta}, \mathbf{x}_i)} f(\mathbf{o}_i, \boldsymbol{\theta}) do_i^K ... d_i^1.$$

The inner loop estimates the Cox model at each iteration until no change occurs in the estimated frailty terms.

B Robustness analysis

B.1 Quits, firings, and other reasons

In this section, we define an alternative classification to job separations based on the reason for job termination: quits, firings, and other reasons. A detailed description of the reasons for job separations is available in Appendix C of the main text.

Our interest in quits partly comes from the intuition that, from a worker's perspective, job spells should be shorter for those jobs created during a recession, because workers are willing to take a low match-quality job foreseeing they will quit to take a better job. Accordingly, it is natural to categorize quits due to reasons other than taking another job as "other reasons." Unfortunately, we can identify quits to take a better job only after the survey in 1990. For the surveys before 1990, this type of separation is implicitly recorded under a general category called "quit for other reasons," or, in some cases, under an even more general category called "other (specify)." This general category corresponds roughly to 40% of our job separations, and classifying them as "other reasons" in our analysis would severely bias our results. To recover some of the quits for taking another job, we include a job spell in the "quits" category if that job ended for one of these "other" reasons but the individual had a job lined up. This procedure results in re-labeling 60% of the "quit for other reasons" as "quits" for our empirical analysis. Several other quit categories appear only in 1979 or after 2002. We apply this procedure to these categories as well. In particular, the survey contained a category called "quit because wages are too low" in 1979 only, and two other categories called "quit because of employment conditions" and "quit because didn't like job, boss, coworkers, pay and benefits" in 2002 and onwards. The motivation to end an employment relationship for these types of separations is similar to quitting to take a better job. In fact, we found that all of the individuals who are in one of these three categories also reported they had jobs lined up. Another category, called "quit to look for a better job," appeared after the survey in 1990. By definition, no job was lined up after this job. We exclude this type of separation from the quits category, because the individual might have been offered a low wage and was forced to quit his job.

Ideally, we would want the "firings" category to include only discharges, firings, and layoffs. For example, temporary and seasonal jobs are set for a fixed term regardless of match quality, and potentially behave differently over the business cycle. Similarly, all jobs, regardless of match quality, are terminated after a plant closure, and we would want to classify this type of job terminations under a separate category. Unfortunately, layoffs were mixed into the same category with plant closures and temporary and seasonal jobs before the survey in 1984, although discharges and firings can be identified for all the survey years. Therefore, we merge all the separations initiated by the firm under the "firings" category. Finally, we also exclude reasons 27 through 30 because they are related to self-employment.

Figure 6 shows the distribution of (completed) job spells by termination reason in our sample used in competing-risk regressions with frailty specification. This figure reflects two features of our sample. First, the reason categories introduced in 2002 are a tiny fraction of the total jobs. Second, about 40% of all the (completed) job spells are due to the "quits for other reasons" category, which is available in all survey years. Of these jobs, three fifths already had a job lined up and were therefore classified as quits.

Of interest to us is how well quits match with EE transitions and firings match with EU transitions. Table 14 shows the cross tabulation of these job-separation categories from our sample with the longer panel. As is evident from the table, we see most of the quits are an

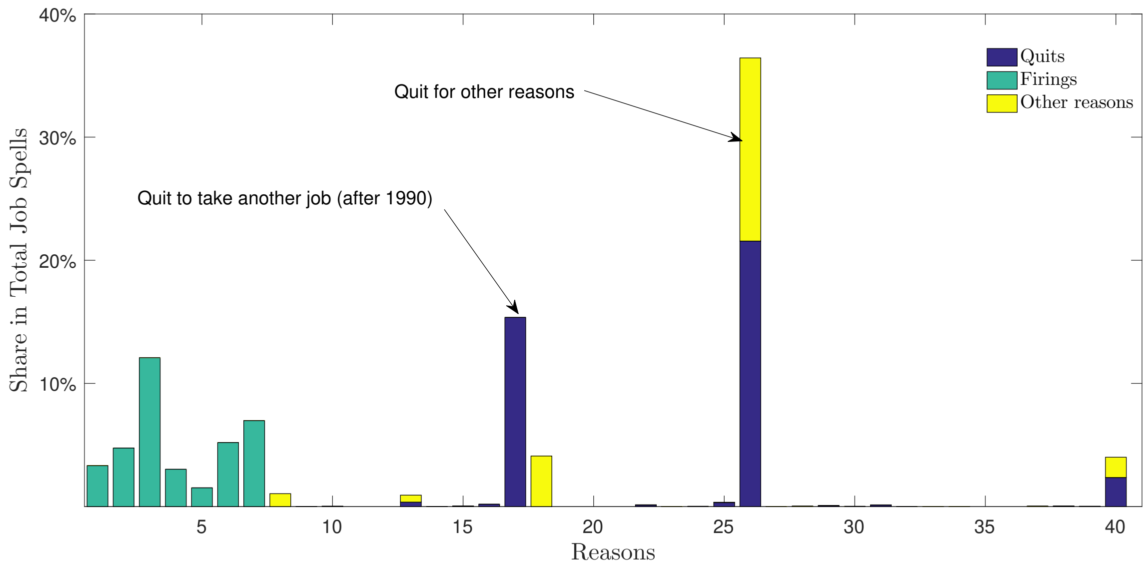

Figure 6: Distribution of job spells by reason.

Transition Type

Reason	EE	EU	EN	Total
Quits	6,288	159	315	6,762
Firings	0	5,075	1,060	6,135
Other reasons	1,452	964	1,310	3,726
Total	7,740	6,198	2,685	16,623

Table 14: Cross Tabulation of Job Separations by Transition Type and Reason: Calculations are from our sample used in competing-risks regressions with frailty specification.

EE transitions and most of the firings are an EU transition. Not surprisingly, our hazard estimations in Tables 15 and 16 are similar to those in Tables 1 and 2, with quits corresponding to EE transitions and firings corresponding to EU transitions. In the subhazard regression, where the coefficients are interpretable, the unemployment rate upon match has a positive effect on quits and a negative effect on firings. The current unemployment rate has a negative effect on quits and a positive effect on firings. The effects on other reasons are mixed, but mostly they are statistically insignificant. Again, these effects of u_t can easily be understood by the procyclical aggregate quit rate and countercyclical aggregate firing rate.

Figures 7, 8, and 9 correspond to Figures 1, 2, and 3 in the main text. The results are very similar, with quits corresponding to EE transitions and firings corresponding to EU transitions.

	Normal Frailty			Stratified		
Variable	Quits	Firings	Other	Quits	Firings	Other
u_0	.053***	−.055***	.012	.047**	−.038*	.018
	(.010)	(.011)	(.015)	(.018)	(.020)	(.026)
u_t	−.106***	.123***	−.054***	−.121***	.094***	−.053*
	(.010)	(.011)	(.015)	(.018)	(.020)	(.028)
SWAGE	−.560***	−.376***	−.961***	−.985***	−.163**	-.998***
	(.033)	(.037)	(.051)	(.077)	(.074)	(.110)
AGE	.014	−.018	−.006	.004	−.075***	−.047
	(.013)	(.013)	(.018)	(.017)	(.021)	(.030)
SQAGE	−.001***	.000	−.000*	−.001*	.001**	.000
	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
UNION	−.352***	.167***	−.135**	−.205***	.114	−.114
	(.045)	(.041)	(.060)	(.077)	(.071)	(.093)
HS	.117**	−.150***	−.313***	-	-	
	(.042)	(.041)	(.050)	-	-	
COL	.497***	−0.450***	−.611***	-	-	
	(.057)	(.069)	(.086)	-	-	
NWHITE	−.124***	.288***	.208***	-	-	
	(.033)	(.033)	(.042)	-	-	
Frailty (Variance)	.398***	.406***	.515***	-	-	
Occurrence:	6,762	6,135	3,726	6,597	5,996	3,628
# of job spells:		19,544			18,747	
# of individuals:		4,293			3,496	
# of right-censored:		2,921			2,526	

Table 15: Cause-Specific Hazard Estimation under Quits, Firing, and Other Reasons Classifications: These estimations use data from all the survey years available. SWAGE=natural logarithm of real starting wages; SQAGE=age squared; UNION=1 if the job is covered under a union contract or collective bargaining agreement; HS=1 if the respondent is a high-school graduate; COL=1 if he completed 16 or more years of education; NWHITE=1 if the respondent is black or Hispanic. Standard errors are given in parentheses. Coefficient estimates for industry and occupation variables are not reported. Significance of the variance of the frailty terms are based on a likelihood-ratio test comparing the model with and without frailty terms. For stratified regressions, standard errors are clustered over individuals. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

Variable	Normal Frailty Quits	Normal Frailty Firings	Normal Frailty Other	Stratified Quits	Stratified Firings	Stratified Other
u_0	.125***	−.106***	.067***	.087***	−.071***	.053**
	(.017)	(.019)	(.028)	(.017)	(.018)	(.025)
u_t	−.217***	.170***	−.138***	−.182***	.140***	−.097***
	(.020)	(.019)	(.032)	(.019)	(.019)	(.028)
SWAGE	−.192***	−.021	−.590***	−.611***	−.290***	-.442***
	(.047)	(.050)	(.070)	(.067)	(.062)	(.086)
AGE	.040**	.017	.004	.058***	−.040**	.001
	(.017)	(.017)	(.024)	(.019)	(.018)	(.027)
SQAGE	−.001***	−.000	−.000*	−.001***	.000*	−.000
	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
UNION	−.326***	.283***	−.082	−.242***	.201***	−.111*
	(.060)	(.051)	(.075)	(.067)	(.062)	(.085)
HS	.249***	−.078	−.285***	-	-	
	(.082)	(.069)	(.080)	-	-	
COL	.687***	−0.446***	−.618***	-	-	
	(.111)	(.112)	(.133)	-	-	
NWHITE	−.216***	.278***	.177***	-	-	
	(.042)	(.041)	(.051)	-	-	
Frailty (Variance)	.301***	.282***	.344***	-	-	
Occurrence:	6,762	6,135	3,726	6,597	5,996	3,628
# of job spells:		19,544			18,747	
# of individuals:		4,293			3,496	
# of right-censored:		2,921			2,526	

Table 16: Subhazard Estimation under Quits, Firing, and Other Reasons Classifications: These estimations use data from all the survey years available. SWAGE=natural logarithm of real starting wages; SQAGE=age squared; UNION=1 if the job is covered under a union contract or collective bargaining agreement; HS=1 if the respondent is a high-school graduate; COL=1 if he completed 16 or more years of education; NWHITE=1 if the respondent is black or Hispanic. Standard errors are given in parentheses. Coefficient estimates for industry and occupation variables are not reported. Significance of the variance of the frailty terms are based on a likelihood-ratio test comparing the model with and without frailty terms. For stratified regressions, standard errors are clustered over individuals and corrected for KM weighting errors. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

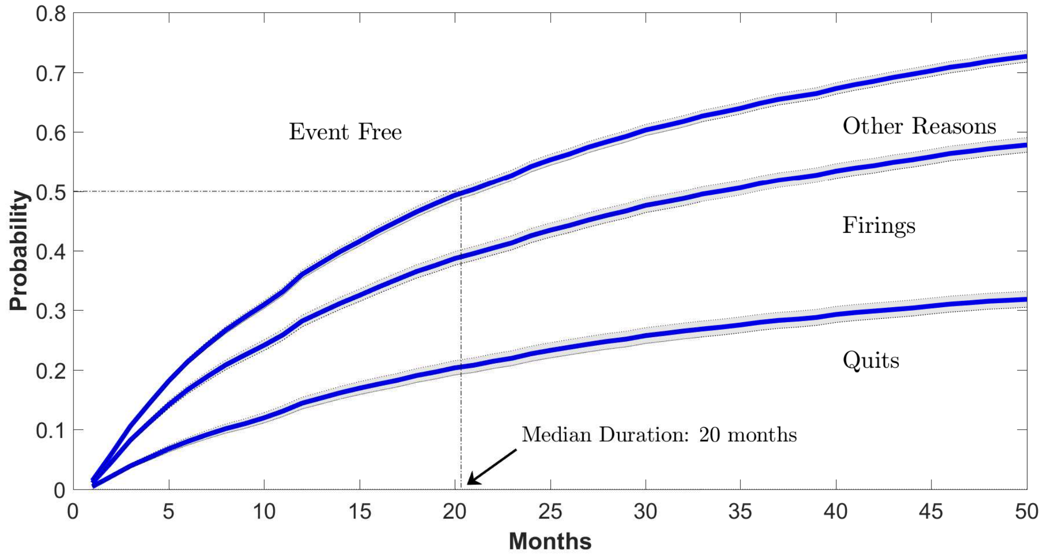

Figure 7: Cumulative Incidence Functions for Quits, Firings, and Other Reasons: The cumulative incidence functions are stacked so that the distance between two curves represents the probabilities of the different events. Shaded areas around each transition type represent 95% confidence intervals.

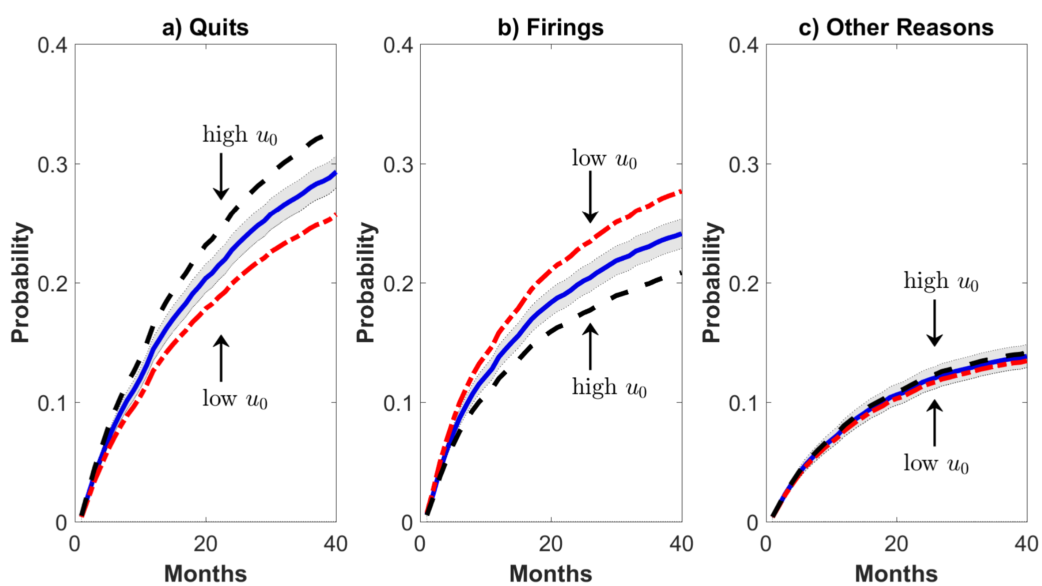

Figure 8: Changes in Cumulative Incidence Functions in Response to a Change in u_0. Shaded areas around each transition type represent 95% confidence intervals.

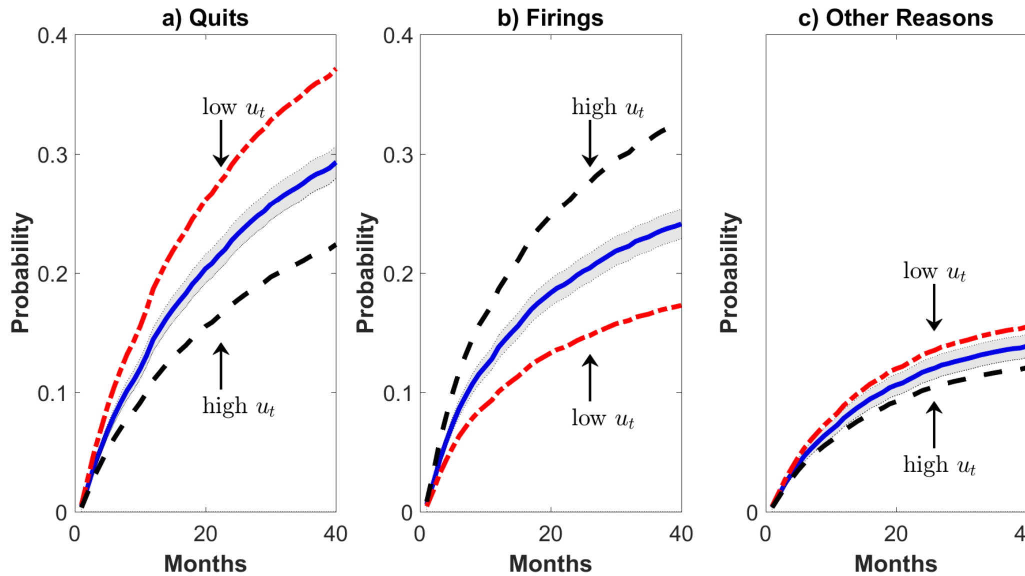

Figure 9: Changes in Cumulative Incidence Functions in Response to a Change in u_t. Shaded areas around each transition type represent 95% confidence intervals.

B.2 Short panel regressions

In this section, we repeat our analysis by dropping the job spells that started before 1988. This time period corresponds to the period analyzed by Bowlus (1995). The job spells that started before 1988 but ended after 1988 are recorded as right censored on the last day of 1988.

Tables 17 and 18 present our results. Our main results for EE and EU transitions are robust to this time restriction. The regression results for EE and EU transitions from the stratified regressions generally agree with our earlier findings with the longer panel, although the estimates are not statistically significant. We note a large decline in our sample size due to the time-period restriction and the large estimated standard errors.

Moreover, distinguishing job separations into unemployment and out of the labor force matters for the jobs held at younger ages. When we pool EU and EN transitions in our regression for the shorter panel, the effects of u_0 and u_t on transition to non-employment are smaller (in absolute value) than the coefficient estimates for EU transitions alone. For u_0, the coefficient estimate is no longer significantly different from 0.[37] This finding is not surprising given that transitions to unemployment and out of the labor force follow different cyclical patterns. We suspect that, compared to our long panel sample, the EN category in our short panel has relatively more transitions out of labor force than to unemployment. Fujita

[37]For cause-specific regressions under frailty, the estimates for u_0 and u_t are $-.027$ and .031 with standard errors of 0.16 and 0.16, respectively.

Variable	Normal Frailty *EE*	*EU*	*EN*	Stratified *EE*	*EU*	*EN*
u_0	.050***	−.045**	.015	.018	−.053	−.021
	(.017)	(.020)	(.029)	(.034)	(.033)	(.056)
u_t	−.187***	.062***	−.056*	−.137***	.077**	.010
	(.018)	(.019)	(.029)	(.035)	(.033)	(.059)
SWAGE	−.811***	−.642***	−.816***	−1.126***	−.483***	−.894***
	(.057)	(.064)	(.101)	(.138)	(.132)	(.232)
AGE	−.069	−.147	−.327**	−.111	−.229	−.536**
	(.083)	(.092)	(.140)	(.143)	(.160)	(.273)
SQAGE	−.002	.003	.006*	−.002	.004	.009
	(.002)	(.002)	(.003)	(.003)	(.004)	(.006)
UNION	−.238***	.192***	.015	−.034	.023	−.016
	(.058)	(.055)	(.091)	(.099)	(.104)	(.154)
HS	.007	−.090	−.259***	-	-	
	(.052)	(.055)	(.083)	-	-	
COL	.204*	−0.724***	−.492**	-	-	
	(.087)	(.124)	(.180)	-	-	
NWHITE	−.162***	.176***	.337***	-	-	
	(.049)	(.050)	(.076)	-	-	
Frailty (Variance)	.422***	.454***	.841***	-	-	
Occurrence:	3,164	2,784	1,180	3,014	2,668	1,083
# of job spells:		9,242			8,306	
# of individuals:		3,142			2,206	
# of right-censored:		2,114			1,541	

Table 17: Cause-Specific Hazard Estimation under EE, EU, and EN Classifications: These estimations use data from 1979 to 1988. SWAGE=natural logarithm of real starting wages; SQAGE=age squared; UNION=1 if the job is covered under a union contract or collective bargaining agreement; HS=1 if the respondent is a high-school graduate; COL=1 if he completed 16 or more years of education; NWHITE=1 if the respondent is black or Hispanic. Standard errors are given in parentheses. Coefficient estimates for industry and occupation variables are not reported. Significance of the variance of the frailty terms are based on a likelihood-ratio test comparing the model with and without frailty terms. For stratified regressions, standard errors are clustered over individuals. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

	Normal Frailty			Stratified		
Variable	*EE*	*EU*	*EN*	*EE*	*EU*	*EN*
u_0	.185***	−.034	.085**	.093***	−.030	.107**
	(.027)	(.030)	(.048)	(.030)	(.031)	(.051)
u_t	−.315***	.079***	−.102**	−.207***	.090***	−.074*
	(.031)	(.031)	(.055)	(.033)	(.031)	(.054)
SWAGE	−.406***	−.291***	−.394***	−.805***	−.021	-.220*
	(.084)	(.089)	(.132)	(.112)	(.112)	(.164)
AGE	.099	−.036	−.274*	.158	−.043	−.309*
	(.115)	(.116)	(.170)	(.133)	(.143)	(.232)
SQAGE	−.002	.000	.004	−.004	−.000	.004
	(.003)	(.003)	(.004)	(.003)	(.003)	(.005)
UNION	−.270***	.240***	.022	−.052	.047	.053
	(.080)	(.069)	(.110)	(.090)	(.092)	(.142)
HS	.069	−.064	−.248**	-	-	
	(.090)	(.082)	(.117)	-	-	
COL	.341**	−.717***	−.471**	-	-	
	(.156)	(.189)	(.265)	-	-	
NWHITE	−.193***	.189***	.356***	-	-	
	(.057)	(.056)	(.085)	-	-	
Frailty (Variance)	.383***	.359***	.683***	-	-	
Occurrence:	3,164	2,784	1,180	3,014	2,668	1,083
# of job spells:		9,242			8,306	
# of individuals:		3,142			2,206	
# of right-censored:		2,114			1,541	

Table 18: Subhazard Estimation under EE, EU, and EN Classifications: These estimations use data from 1979 to 1988. SWAGE=natural logarithm of real starting wages; SQAGE=age squared; UNION=1 if the job is covered under a union contract or collective bargaining agreement; HS=1 if the respondent is a high-school graduate; COL=1 if he completed 16 or more years of education; NWHITE=1 if the respondent is black or Hispanic. Standard errors are given in parentheses. Coefficient estimates for industry and occupation variables are not reported. Significance of the variance of the frailty terms are based on a likelihood-ratio test comparing the model with and without frailty terms. For stratified regressions, standard errors are clustered over individuals and corrected for KM weighting errors. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

and Ramey (2006) show that the cyclical movements in job separations to unemployment and out of the labor force for young individuals are roughly equal to each other while the former dominates for prime-age individuals. Our results are broadly consistent with their findings.

B.3 Local labor market conditions and the role of unemployment benefits

In each survey round, individuals report the state of residence at the time of the interview. This information is available from the restricted access geo-code data. We use this information to add some state-level variables to our analysis. One potential measurement issue with this variable is that it indicates the location of the individual at the time of the interview, which may be different than the location of the job if the job has already ended. Therefore, we created two samples. In the first one, we include only the current jobs for which we are certain about the location of the job. In the second sample, we assume the location of the job is the same as the location of the individual in the survey year that is closest to the start date of that job.

Under these specifications, we made the following changes to our sample. First, when creating the variables u_0 and u_t, we replaced the national-level unemployment rate with the state-level unemployment rate obtained from BLS. Second, the unemployment insurance (UI) policies vary significantly across states, and these benefits have a potential impact on individual's employment decisions. To control for these effects, we obtained the state-level UI replacement-rate series from the Department of Labor and included it in our regressions. These series come from the UI Financial Data Handbook and are available annually from 1979.

Our estimation results from the regressions using the sample with the current jobs only are presented in Tables 19 and 20. Tables 21 and 22 show our results using all the job spells with the imputed state of residence information. Overall, the effects of u_0 on EE and EU transitions are smaller, but they are still statistically significant and their signs agree with our results in the main text. These findings suggest some of the effects we measure in the main analysis reflect variation across states and the local labor-market policies.

B.4 Non-parametric estimation of the frailty model

In this section, we relax the normality and independence assumptions we maintained in the cause-specific regressions with frailty. Although Heckman and Honoré (1989) proves identification of the competing-risk model with frailty under fairly general conditions, they

Variable	Normal Frailty *EE*	Normal Frailty *EU*	Normal Frailty *EN*	Stratified *EE*	Stratified *EU*	Stratified *EN*
u_0^s	.034***	−.037***	−.050***	.033**	−.041**	.022
	(.008)	(.009)	(.015)	(.015)	(.018)	(.030)
u_t^s	−.079***	.116***	.026*	−.093***	.106***	−.052*
	(.009)	(.009)	(.014)	(.016)	(.033)	(.030)
UIRR	−.860**	.602	−1.364**	−.065	2.587**	1.534
	(.290)	(.346)	(.503)	(.752)	(1.031)	(1.892)
SWAGE	−.540***	−.380***	−.709***	−.968***	−.116	−.733***
	(.036)	(.044)	(.066)	(.087)	(.092)	(.232)
AGE	−.003	.003	−.026	−.025	−.058**	−.152***
	(.014)	(.016)	(.024)	(.024)	(.026)	(.046)
SQAGE	−.001***	−.000	.000	−.000	.000	.002**
	(.000)	(.000)	(.000)	(.000)	(.000)	(.001)
UNION	−.377***	.122**	−.066	−.194**	.101	.101
	(.049)	(.050)	(.079)	(.083)	(.088)	(.140)
HS	.069	−.226***	−.311***	-	-	
	(.046)	(.050)	(.070)	-	-	
COL	.340***	−.560***	−.644***	-	-	
	(.063)	(.080)	(.116)	-	-	
NWHITE	−.124***	.307***	.347***	-	-	
	(.036)	(.041)	(.058)	-	-	
Frailty (Variance)	.429***	.413***	.725***	-	-	
Occurrence:	5,666	4,168	2,031	5,545	4,101	1,958
# of job spells:		14,712			14,109	
# of individuals:		4,121			3,518	
# of right-censored:		2,847			2,505	

Table 19: Cause-Specific Hazard Estimation under EE, EU, and EN Classifications with State Related Variables: These estimations use data from all the survey years available. The sample includes only the ongoing jobs at the time of the interview. u_0^s and u_t^s represent the state-level unemployment rate at the start of the job spell and at duration t, respectively. UIRR=state-level UI benefit replacement rate; SWAGE=natural logarithm of real starting wages; SQAGE=age squared; UNION=1 if the job is covered under a union contract or collective bargaining agreement; HS=1 if the respondent is a high-school graduate; COL=1 if he completed 16 or more years of education; NWHITE=1 if the respondent is black or Hispanic. Standard errors are given in parentheses. Coefficient estimates for industry and occupation variables are not reported. Significance of the variance of the frailty terms are based on a likelihood-ratio test comparing the model with and without frailty terms. For stratified regressions, standard errors are clustered over individuals. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

	Normal Frailty			Stratified		
Variable	*EE*	*EU*	*EN*	*EE*	*EU*	*EN*
u_0^s	.075***	−.057***	−.049**	.055***	−.050***	−.011
	(.013)	(.015)	(.024)	(.014)	(.017)	(.025)
u_t^s	−.145***	.140***	.013	−.127***	.133***	−.032*
	(.014)	(.015)	(.025)	(.015)	(.017)	(.025)
UIRR	−.752*	.947**	−1.113*	−.721	2.222***	−.018
	(.467)	(.517)	(.731)	(.717)	(.931)	(1.211)
SWAGE	−.266***	−.075	−.367***	−.760***	−.277***	−.117
	(.052)	(.060)	(.085)	(.073)	(.075)	(.109)
AGE	.040**	.036**	.012	-.042**	−.017	−.076**
	(.019)	(.020)	(.029)	(.022)	(.024)	(.038)
SQAGE	−.001***	−.001**	−.000	−.001***	.000	.001**
	(.000)	(.000)	(.000)	(.000)	(.000)	(.001)
UNION	−.350***	.264***	.029	−.263**	.193***	−.003
	(.065)	(.061)	(.093)	(.074)	(.076)	(.110)
HS	.206**	−.152**	−.229**	-	-	
	(.086)	(.078)	(.102)	-	-	
COL	.530***	−.514***	−.587***	-	-	
	(.114)	(.123)	(.170)	-	-	
NWHITE	−.236***	.294***	.315***	-	-	
	(.049)	(.048)	(.067)	-	-	
Frailty (Variance)	.300***	.197***	.520***	-	-	
Occurrence:	5,666	4,168	2,031	5,545	4,101	1,958
# of job spells:		14,712			14,109	
# of individuals:		4,121			3,518	
# of right-censored:		2,847			2,505	

Table 20: Subhazard Estimation under EE, EU, and EN Classifications with State Related Variables: These estimations use data from all the survey years available. The sample includes only the ongoing jobs at the time of the interview. u_0^s and u_t^s represent the state-level unemployment rate at the start of the job spell and at duration t, respectively. UIRR=state-level UI benefit replacement rate; SWAGE=natural logarithm of real starting wages; SQAGE=age squared; UNION=1 if the job is covered under a union contract or collective bargaining agreement; HS=1 if the respondent is a high-school graduate; COL=1 if he completed 16 or more years of education; NWHITE=1 if the respondent is black or Hispanic. Standard errors are given in parentheses. Coefficient estimates for industry and occupation variables are not reported. Significance of the variance of the frailty terms are based on a likelihood-ratio test comparing the model with and without frailty terms. For stratified regressions, standard errors are clustered over individuals and corrected for KM weighting errors. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

	Normal Frailty			Stratified		
Variable	*EE*	*EU*	*EN*	*EE*	*EU*	*EN*
u_0	.030***	−.038***	−.051***	.028**	−.036**	−.004
	(.008)	(.008)	(.013)	(.013)	(.015)	(.023)
u_t	−.086***	.110***	.042***	−.096***	.097***	−.014
	(.008)	(.008)	(.013)	(.014)	(.016)	(.025)
UIRR	−.974***	.291	−.796*	−.552	1.216*	.659
	(.250)	(.284)	(.423)	(.626)	(.703)	(1.111)
SWAGE	−.603***	−.422***	−.754***	−.977***	−.253***	−.657***
	(.031)	(.036)	(.056)	(.069)	(.073)	(.114)
AGE	−.005	−.035***	−.082***	−.020	−.081***	−.191***
	(.012)	(.012)	(.140)	(.020)	(.019)	(.033)
SQAGE	−.000	.000	.001***	−.000	.001***	.002***
	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
UNION	−.311***	.114***	−.097	−.167**	.069	−.037
	(.042)	(.041)	(.068)	(.070)	(.069)	(.095)
HS	.054	−.150***	−.309***	-	-	
	(.039)	(.041)	(.058)	-	-	
COL	.380***	−.407***	−.615**	-	-	
	(.053)	(.066)	(.100)	-	-	
NWHITE	−.155***	.258***	.410***	-	-	
	(.031)	(.033)	(.048)	-	-	
Frailty (Variance)	.357***	.395***	.698***	-	-	
Occurrence:	7,804	6,460	3,004	7,659	6,351	2,899
# of job spells:		20,124			19,422	
# of individuals:		4,279			3,577	
# of right-censored:		2,856			2,513	

Table 21: Cause-Specific Hazard Estimation under *EE*, *EU*, and *EN* Classifications with State Related Variables: These estimations use data from all the survey years available. The sample includes all the job spells. u_0^s and u_t^s represent the state-level unemployment rate at the start of the job spell and at duration t, respectively. UIRR=state-level UI benefit replacement rate; SWAGE=natural logarithm of real starting wages; SQAGE=age squared; UNION=1 if the job is covered under a union contract or collective bargaining agreement; HS=1 if the respondent is a high-school graduate; COL=1 if he completed 16 or more years of education; NWHITE=1 if the respondent is black or Hispanic. Standard errors are given in parentheses. Coefficient estimates for industry and occupation variables are not reported. Significance of the variance of the frailty terms are based on a likelihood-ratio test comparing the model with and without frailty terms. For stratified regressions, standard errors are clustered over individuals. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

	Normal Frailty			Stratified		
Variable	*EE*	*EU*	*EN*	*EE*	*EU*	*EN*
u_0	.094***	−.063***	−.051**	.066***	−.045***	−.013
	(.013)	(.015)	(.024)	(.013)	(.015)	(.023)
u_t	−.175***	.140***	.036*	−.154***	.124***	−.001
	(.015)	(.015)	(.025)	(.014)	(.015)	(.024)
UIRR	−.860**	.749**	−.459	−1.098**	1.541***	.514
	(.412)	(.447)	(.633)	(.601)	(.655)	(.928)
SWAGE	−.255***	−.088**	−.328***	−.668***	.173***	−.074
	(.045)	(.050)	(.072)	(.060)	(.062)	(.085)
AGE	.047***	.006	−.030*	.056***	−.029*	−.095***
	(.016)	(.016)	(.023)	(.019)	(.018)	(.029)
SQAGE	−.001***	−.000	.000*	−.001**	.000	.001***
	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
UNION	−.296***	.237***	−.046	−.214***	.142***	−.070
	(.056)	(.052)	(.079)	(.061)	(.060)	(.086)
HS	.173**	−.082	−.258***	-	-	
	(.076)	(.070)	(.092)	-	-	
COL	.526***	−.392***	−.623***	-	-	
	(.103)	(.109)	(.156)	-	-	
NWHITE	−.255***	.248***	.394***	-	-	
	(.040)	(.043)	(.060)	-	-	
Frailty (Variance)	.259***	.283***	.506***	-	-	
Occurrence:	7,804	6,460	3,004	7,659	6,351	2,899
# of job spells:		20,124			19,422	
# of individuals:		4,279			3,577	
# of right-censored:		2,856			2,513	

Table 22: Subhazard Estimation under EE, EU, and EN Classifications with State Related Variables: These estimations use data from all the survey years available. The sample includes all the job spells. u_0^s and u_t^s represent the state-level unemployment rate at the start of the job spell and at duration t, respectively. UIRR=state-level UI benefit replacement rate; SWAGE=natural logarithm of real starting wages; SQAGE=age squared; UNION=1 if the job is covered under a union contract or collective bargaining agreement; HS=1 if the respondent is a high-school graduate; COL=1 if he completed 16 or more years of education; NWHITE=1 if the respondent is black or Hispanic. Standard errors are given in parentheses. Coefficient estimates for industry and occupation variables are not reported. Significance of the variance of the frailty terms are based on a likelihood-ratio test comparing the model with and without frailty terms. For stratified regressions, standard errors are clustered over individuals and corrected for KM weighting errors. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

Variable	EE	EU	EN
Variance	.164	.175	.476
Skewness	−.104	−.092	−.513
Correlation Matrix			
EE	1.000	–	–
EU	.300	1.000	–
EN	.328	.324	1.000

Table 23: Non-Parametric Estimation of the Frailty Distribution

do not provide a method for the estimation. We implemented the general version of the EM algorithm described in Section A.3 of this online Appendix. Some of the practical issues are as follows. We consider the unknown distribution of the frailty terms as a discrete grid defined in R^3. We choose 11 points for each dimension and obtain a mesh grid of these points so that we have 11^3 points in total over this three-dimensional rectangular grid.[38] The outer loop in the EM algorithm estimates the probability mass function over this rectangular grid. For each dimension, we choose a symmetric grid around zero such that the end points are two standard deviations away from zero. We use the standard-deviation estimates from the regressions under normality and independence assumptions. We also normalize the probabilities so that the mean of each frailty term is equal to 0. The problem is fairly high dimensional and potentially has multiple local maxima. To ensure we obtained the global maximum, we searched for the solution multiple times by starting from a different set of initial points. We consistently get results that are comparable at machine precision.

In Table 23, we report the variance and skewness of the estimated distribution followed by the correlation matrix. The results are fairly at odds with the normality and independence assumptions: the frailty terms are skewed left and positively correlated with each other. Nonetheless, our parameter estimates in Table 24 are very close to those obtained under normality and independence assumptions.

B.5 Interaction of unemployment rate with industry

We included industry fixed effects in the main text. In this section, we further investigate whether the estimated effects of u_0 and u_t vary over industries. Given the small sample size, we defined even broader categories for industries, namely, goods-producing industries and service-providing industries. Then, we included the interaction of the dummy variables for

[38] Although the number of mass points we choose seems arbitrary, van den Berg (2001) notes it is rare to find more than a few mass points in the non-parametric estimation of the proportional hazard models without competing risks.

Variable	EE	EU	EN
u_0	.056***	−.054***	−.046***
	(.009)	(.009)	(.017)
u_t	−.103***	.117***	.033***
	(.010)	(.008)	(.016)
SWAGE	−.494***	−.499***	−.796***
	(.027)	(.029)	(.055)
AGE	−.014	−.041***	−.104***
	(.011)	(.011)	(.019)
SQAGE	−.000*	.000**	.001***
	(.000)	(.000)	(.000)
UNION	−.330***	.126***	−.100***
	(.039)	(.033)	(.067)
HS	.016	−.126***	−.282***
	(.034)	(.031)	(.057)
COL	.302***	−0.311***	−.476***
	(.048)	(.052)	(.099)
NWHITE	−.165***	.220***	.384***
	(.029)	(.026)	(.047)
Occurrence:	8,063	6,657	3,100
# of job spells:		20,741	
# of individuals:		4,331	
# of right-censored:		2,921	

Table 24: Non-parametric Estimation of the Frailty Model: These estimations use data from all the survey years available. SWAGE=natural logarithm of real starting wages; SQAGE=age squared; UNION=1 if the job is covered under a union contract or collective bargaining agreement; HS=1 if the respondent is a high-school graduate; COL=1 if he completed 16 or more years of education; NWHITE=1 if the respondent is black or Hispanic. Standard errors are given in parentheses. Coefficient estimates for industry and occupation variables are not reported. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

these industries with u_0 and u_t along with the original industry dummy variables. Our results are presented in Tables 25 and 26. One notable difference from the main text is that the effect of u_0 on EU transitions for goods producing industries is still negative but insignificant. The point estimates are also smaller than those for the service providing industries by a great margin.

	Normal Frailty			**Stratified**		
Variable	EE	EU	EN	EE	EU	EN
u_0 (Goods)	.060***	−.024	−.037	.068***	−.009	.012
	(.015)	(.015)	(.024)	(.022)	(.025)	(.044)
u_0 (Services)	.056***	−.093***	−.065***	.054**	−.101***	−.031
	(.013)	(.016)	(.023)	(.022)	(.027)	(.036)
u_t (Goods)	−.139***	.121***	.041*	−.153***	.088***	−.013
	(.015)	(.014)	(.023)	(.068)	(.026)	(.046)
u_t (Services)	−.092***	.123***	.036*	−.103***	.129***	−.046
	(.012)	(.014)	(.021)	(.017)	(.026)	(.037)
Occurrence:	8,063	6,657	3,100	7,913	6,542	2,991
# of job spells:		20,741			19,998	
# of individuals:		4,331			3,588	
# of right-censored:		2,921			2,552	

Table 25: Cause-Specific Hazard Estimation under EE, EU, and EN Classifications with Industry Interaction Variables: These estimations use data from all the survey years available. Although not reported, these regressions include all the other explanatory variables from the main analysis. Standard errors are given in parentheses. For stratified regressions, standard errors are clustered over individuals. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

	Normal Frailty			Stratified		
Variable	*EE*	*EU*	*EN*	*EE*	*EU*	*EN*
u_0 (Goods)	.144***	−.061**	−.052	.115***	−.029	−.018
	(.024)	(.025)	(.043)	(.026)	(.024)	(.041)
u_0 (Services)	.127***	−.150***	−.073**	.091***	−.127***	−.033
	(.020)	(.026)	(.041)	(.021)	(.026)	(.038)
u_t (Goods)	−.278***	.161***	.028	−.242***	.135***	.017
	(.029)	(.026)	(.046)	(.028)	(.025)	(.044)
u_t (Services)	−.186***	.178***	.035	−.152***	.166***	−.011
	(.023)	(.025)	(.042)	(.022)	(.024)	(.039)
Occurrence:	8,063	6,657	3,100	7,913	6,542	2,991
# of job spells:		20,741			19,998	
# of individuals:		4,331			3,588	
# of right-censored:		2,921			2,552	

Table 26: Subazard Estimation under EE, EU, and EN Classifications with Industry Interaction Variables: These estimations use data from all the survey years available. Although not reported, these regressions include all the other explanatory variables from the main analysis. Standard errors are given in parentheses. For stratified regressions, standard errors are clustered over individuals. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively.

Additional References for the online Appendix

[1] Therneau, T. M., Grambsch, P. M., and Pankratz, V. S. (2003): "Penalized Survival Models and Frailty," *Journal of Computational and Graphical Statistics* 12, 156–175.

[2] Van den Berg, Gerard J. (2001): "Duration Models: Specification, Identification and Multiple Durations," *Handbook of Econometrics* 5, 3381–3460.