The Churn Ladder

Lawrence Warren David Wiczer*

March 27, 2026

Abstract

In job-ladder models, wage dispersion arises because firms trade off higher wages against the risk of losing workers to competitors. We validate the empirical relevance of a "churn ladder"—a systematic pattern in which workers move from low-wage, high-turnover firms to high-wage, low-turnover firms through job-to-job transitions. The canonical model, however, cannot generate the steep churn gradient we observe in the data: poaching alone produces too little cross-firm dispersion in turnover. What it misses is that lower-wage firms also lose more workers to non-employment, substantially amplifying churn differences across the ladder. We introduce heterogeneity in both firm and worker separation rates, which generates endogenous sorting and allows the model to match observed joint patterns of wages and churn. Worker-level heterogeneity alone accounts for half the variation in firm-level churn. Ignoring wage-dependent separations biases estimates of firms' separation elasticities and overstates monopsony power.

*U.S. Census Bureau, CES; Federal Reserve Bank of New York and IZA. We would like to thank Aniket Baksy, Serdar Birinci, Arin Dube, Eva Carceles-Poveda, Nik Engbom, Ryan Michaels, Chris Moser, Ioannis Kospentaris and Todd Schoellman and seminar participants at San Francisco Fed, St. Louis Fed, University at Albany, University of North Carolina, Chapel Hill, Rice University, and Stony Brook University for helpful comments. This research uses data from the Census Bureau's Longitudinal Employer Household Dynamics Program, which was partially supported by the following National Science Foundation Grants SES-9978093, SES-0339191 and ITR-0427889; National Institute on Aging Grant AG018854; and grants from the Alfred P. Sloan Foundation. Any opinions and conclusions expressed herein are those of the authors and do not represent the views of the U.S. Census Bureau, the Federal Reserve Bank of New York or the Federal Reserve System. The Census Bureau has ensured appropriate access and use of confidential data and has reviewed these results for disclosure avoidance protection (Federal Statistical Research Data Center Project Number 1819: CBDRB-FY23-P1819-R10342, CBDRB-FY23-P1819-R10598, CBDRB- FY24-P1819-R11361). All errors are our own.

1 Introduction

Most firms do not change size in most periods yet still hire and lose significant numbers of workers; this phenomenon, where gross flows exceed the net change, is known as 'churn'. This worker churn is fundamental in many theories of firm wage setting such as the model of on-the-job search in Burdett and Mortensen (1998). Firms paying a higher wage trade off a lower flow profit for an increase in the expected duration of their matches and a lower rate of churn. This mechanism generates heterogeneity across otherwise identical firms, forming the backbone of a broad class of "job ladder" models that explain residual wage dispersion in the presence of labor search frictions.

In this paper, we combine new evidence from the Census Bureau's Longitudinal Employer-Household Dynamics dataset on the empirical shape of the job ladder in churn and wages. We find significant dispersion in churn over the cross-section of firms and, consistent with a broad set of job ladder models, a negative relationship between churn and firms' wages. We also follow this wage and turnover ladder at the worker level, across their consecutive job transitions.[1] This controls for worker heterogeneity and potential selection, but the churn ladder is still apparent. As workers move from job to job, they move to higher-wage and lower-churn firms, climbing the wage ladder and descending the churn ladder.

Much like earlier work by Hornstein et al. (2011) on the residual wage, we find that the mechanisms inherent in frictional labor market models do not necessarily generate sufficient dispersion observed in the data. Estimating the slope of the wage and churn ladder gives a quantitative benchmark that can be directly compared to calibrated models of on-the-job search. We compute and estimate a version of the Burdett and Mortensen (1998) model that incorporates the wage-retention tradeoff underlying wage and churn ladders, allowing for firm-level dispersion in turnover through on-the-job search. Crucially, on-the-job search itself does not generate enough dispersion in churn to match the data. To match the high turnover rates observed at the bottom rungs of the ladder, there must be high separations

[1] This continuous employment spell with potentially multiple jobs was coined an "employment cycle" in Wolpin (1992). In this paper, we refer to these "employment cycles" simply as "employment spells" and "continuous employment spells," interchangeably.

to non-employment in addition to the elevated poaching rates implied by on-the-job search.

Our model incorporates two possible mechanisms to explain the observed dispersion in churn and the steepness of the churn ladder beyond what poaching alone can generate. First, we allow firms to set policies that make all workers more or less likely to separate from the job into nonemployment, conditional on wages. Second, we introduce latent worker heterogeneity in the probability of separating into nonemployment. This leads to workers differing in their expected employment spell lengths and resulting upward mobility from climbing the job ladder. This gives rise to endogenous sorting: workers with low separation risk are concentrated at high-wage, low-poaching firms, further lowering churn rates at the top of the ladder. In our calibrated model, roughly half of the churn ladder's steepness over the wage distribution comes from worker heterogeneity, about one-third from firms, and the remainder, less than one-fifth, from the poaching mechanism embedded in on-the-job search.

The two explanations, firm policies and worker heterogeneity, also have implications that we can test empirically. We devise two test statistics using the means and variances of tenure conditional on firm-level wages. The intuition is that if firm policies alone drove the churn ladder, separation rates would be constant conditional on wage, and therefore tenure would be exponentially distributed. Worker heterogeneity increases average tenure and variance within the wage bin. We test against the null hypothesis that there is no worker-level heterogeneity in separation probabilities and reject it with both tests. This aligns with the findings in our calibrated model, which gives a role to both worker heterogeneity and firm policies in generating the cross-sectional dispersion in churn.

Our work is related to established models of the job ladder such as Burdett and Mortensen (1998) and Moscarini and Postel-Vinay (2017) which introduce the firm's retention motive into wage determination. In these models, a job ladder can exist both within ex-ante homogeneous firms and across heterogeneous levels of productivity.[2] Theory establishes a clear link between wage rungs and churn rungs of a job ladder, and we explore the empirical magnitude of this effect and how a model of the job ladder can deliver it.

[2] The latter somewhat weakens the strict one-to-one relationship between wage and turnover but maintains the general correlation because wages are still correlated with productivity, which itself has a one-to-one relationship with turnover.

Understanding this relationship is also critical for measuring labor market power. Manning (2003) links the poaching threat faced by the firm to the elasticity of separations with respect to wages and the implied wage markdown in a Burdett and Mortensen (1998) model. Bassier et al. (2021) build on this intuition to causally estimate the labor supply elasticity facing a single firm, focusing on separations to other employers. We show that this measurement exercise often misses an important component of the relevant separation elasticity: separations to nonemployment. In a job ladder environment, the labor supply elasticity facing the firm depends on how wages affect all margins of worker retention, not just job-to-job moves. If higher wages also reduce exits to nonemployment, then approaches that rely only on poaching flows will understate the separation elasticity and therefore overstate firms' monopsony power through an excessively large implied markdown. Incorporating wage-dependent separations to nonemployment raises the measured elasticity and can bring it closer to empirical estimates from dynamic monopsony frameworks and quasi-experimental designs.

In terms of measuring the job ladder in labor market data, seminal work by Topel and Ward (1992) highlighted the importance of job-to-job transitions to workers' wage gains. Moscarini and Postel-Vinay (2008, 2012) first documented a link between job transitions and firm characteristics in the existence of a job ladder in firm size. We extend this measurement of the ladder to firm dynamics, via the churn rate, closely related to Bagger and Lentz (2018) and Sorkin (2018).

We then combine the structure of a partial equilibrium model of on-the-job search along with detailed microdata to back out the distribution of wage offers and resulting worker flows and turnover. Baksy et al. (2024) use a similar approach with wages from the CPS to measure the long-run decline of job transition rates and their cause. Gottfries and Teulings (2023) use match duration and employment histories in NLSY data to measure the contribution of on-the-job search in wage growth and dispersion throughout an "employment cycle."[3] With matched employer-employee data, we can study the interaction of the firm distribution of

[3]Rather than use the wage distribution, they infer the offer distribution from order statistics and the sequence of wage records, expanding on Barlevy (2008).

wages and churn and the worker's job ladder.

Also quite related in incorporating both worker and firm heterogeneity, recent work by Karahan et al. (2019), Lentz et al. (2023), and Lamadon et al. (2024) identifies latent heterogeneity in firms and workers, and the sorting between them. They use mobility for identification but focus primarily on the importance of this heterogeneity to wages, which is complementary to our focus on the distribution of turnover patterns across firms.

Understanding the source of separation heterogeneity has important implications for the role of the job ladder. We find evidence for the firm-specific differences in separation rates that are crucial to the premise of a "slippery" job ladder. This mechanism has been proposed to explain persistent earnings losses after job displacement via repeated unemployment risk in Pinheiro and Visschers (2015) and Jarosch (2023). However, we also reject that it is the only source of heterogeneity in separation rates. Worker-level latent heterogeneity and endogenous sorting produces a meaningful distinction between wage and earnings inequality and their sources. It is also consistent with the findings in Borovičková and Macaluso (2023) that workers with high and low end-of-career earnings experience similar numbers of job transitions, but of differing quality in terms of earnings gain. This paper and our findings also align with a growing literature on latent heterogeneity across workers in their flow patterns. For example, Hall and Kudlyak (2019); Gregory et al. (2021), and Ahn et al. (2023). While these papers primarily focused on the dynamics of worker flows in and out of employment, we find these insights about heterogeneity in workers' separation rates to extend to the distribution of employed workers. This latent worker heterogeneity is also critical in replicating the distribution of their employers, both in their wages and their gross flows.

2 Empirical estimates of wage and job ladders

2.1 Data

This paper's data comes from the LEHD, a matched employer-employee panel consisting of administrative data collected from quarterly state unemployment insurance files. These files cover most of wage and salary employment in the United States with the exception of federal employees, the self-employed, gig workers, and others not typically covered by the unemployment insurance system. Our sample consists of a random 15% sample of workers from a 17 state sample over the time period spanning 1998-2013.[4]

At the employer, we observe firm identifiers at the State - Employer Identification Number (SEIN) level, through which we can link and observe employment and earnings of workers at the firm in that state.[5] We use several measures of establishment characteristics and dynamics from the Quarterly Workforce Indicators (QWI) files. We use employment weights to aggregate establishments within the state to the SEIN level. For each firm, we compute hires as employer-level "accessions," which includes all new employees and does not exclude recalls. We exclude firms that start up or shut down in that quarter as well as those with more than 200% turnover within a quarter.

From our 15% sample of workers, we use the job history file (JHF), which lists earnings at each job spell during a worker's lifetime. For each period in which the worker has earnings, we follow Hahn et al. (2017) by selecting a "dominant" employer as the job with maximum earnings over two quarters. A job-to-job transition is measured as a change in dominant job from one period to the next when the new dominant job had positive earnings in the same quarter as the old dominant job. This measure will potentially misclassify a worker as having made a job transition when they have worked continually at two jobs where the dominant job fluctuates between the two employers. Similarly, this measure also misses some

[4] Our sample includes data from California, Colorado, Hawaii, Idaho, Illinois, Indiana, Kansas, Maine, Maryland, Missouri, Montana, Nevada, North Dakota, Tennessee, Texas, Virginia and Washington. These states represent 42% of national employment.

[5] Multi-state firms will be counted as separate firms in each state. The firm-level data covers the firms' entire workforce, not just our 15% sample.

true job-to-job transitions that occur over a weekend at the seam of two quarters since we require overlapping earnings. Time aggregation also implies that some job-to-job transitions we measure are transitions into non-employment for up to 11 weeks before transitioning to a new job.

We restrict our job transition sample in two principal ways. On the worker side, we restrict to workers who are between 25-65 years old when their job begins. This truncation by age is consistent with much of the earnings risk literature to reduce the prevalence of very short job spells. On the firm side, we remove transitions involving firms with fewer than 10 employees to eliminate large percentage changes from discrete employment changes. We accumulate worker earnings and firm growth and turnover into annual measures before and after a transition. Our sample restrictions result in 1.2 million job-to-job transitions. [6] We deflate earnings in all quarters to 2009 dollars.

2.2 Definitions

Our primary focus is on measuring a firm's "churn," the gross flows above and beyond net employment growth. To measure churn, we measure excess hires, $F_{i,t}$ as the number of hires beyond the ones needed for firm i's net employment growth in that period. This is equivalent to the churn measure in Elsby et al. (2017):

$$\begin{aligned} F_{i,t} =& H_{i,t} - \max(\Delta L_{i,t}, 0) = \min(H_{i,t}, S_{i,t}) \\ =& \frac{1}{2}\left(H_{i,t} - \max(\Delta L_{i,t}, 0) + S_{i,t} - \min(\Delta L_{i,t}, 0)\right) \ . \end{aligned} \tag{2.1}$$

where $H_{\ell,t} = \frac{\sum_{k=1}^{4} \mathbb{H}_{\ell,t-k}}{\frac{1}{2}(L_{\ell,t}+L_{\ell,t-4})}$, $H_{d,t} = \frac{\sum_{k=1}^{4} \mathbb{H}_{d,t+k}}{\frac{1}{2}(L_{d,t}+L_{d,t+4})}$ are the hiring rates and $S_{\ell,t} = \frac{\sum_{k=1}^{4} \mathbb{S}_{\ell,t-k}}{\frac{1}{2}(L_{\ell,t}+L_{\ell,t-4})}$ and $S_{d,t} = \frac{\sum_{k=1}^{4} \mathbb{S}_{d,t+k}}{\frac{1}{2}(L_{d,t}+L_{d,t+4})}$ are the separation rates.

Note that ℓ indexes the origin firm and d the destination firm associated with a transition.[7]

[6] In order to measure annual earnings growth and firm dynamics, we require the origin and destination match to last 6 quarters. This restriction drops a substantial number of observations which is consistent with Hyatt and Spletzer (2017), who report that about $\frac{1}{3}$ of all job matches last less than 1 quarter.

[7] Since these annual measures are chosen to go with a transition at time t, we suppress the t notation for

For a transition in period t, we look 4 quarters prior to the transition and 4 quarters after the transition. So that our measure is not mechanically affected by the transition of the worker we observe, firm characteristics indexed t skip the period of transition, i.e. $t+1$ to $t+4$ and $t-1$ to $t-4$ just like we did with earnings growth.

Note that net job flows, or net employment growth at the firm can also be represented as a function of employment, hires and separations:

$$\Delta L_{\ell,t} = \frac{\sum_{k=1}^{4} \mathbb{H}_{\ell,t-k} - \sum_{k=1}^{4} \mathbb{S}_{\ell,t-k}}{\frac{1}{2}(L_{\ell,t} + L_{\ell,t-4})}, \quad \Delta L_{d,t} = \frac{\sum_{k=1}^{4} \mathbb{H}_{d,t+k} - \sum_{k=1}^{4} \mathbb{S}_{d,t+k}}{\frac{1}{2}(L_{d,t} + L_{d,t+4})}. \tag{2.2}$$

We summarize average earnings of workers at the firm as the log of the average real earnings a firm pays its employees.

3 Empirical evidence on the relationship between churn and wages

We first want to consider the relationship between earnings and churn at the firm level. Figure 1 plots the quarterly excess hire rate, or churn rate, of firms over their percentile rank in terms of average wages. We can see that there is a strong monotonic decline in churn over the distribution of firms' average wages. Churn rates are very high at low-wage firms and at the top of the firm wage distribution, the churn rate is nearly zero. This lines up well with the intuition behind most job ladder models, that firms face a tradeoff between their wage policy and worker retention. It is also consistent with the idea that high-wage firms poach from firms with lower wages.

the firm and just denote them with d and ℓ.

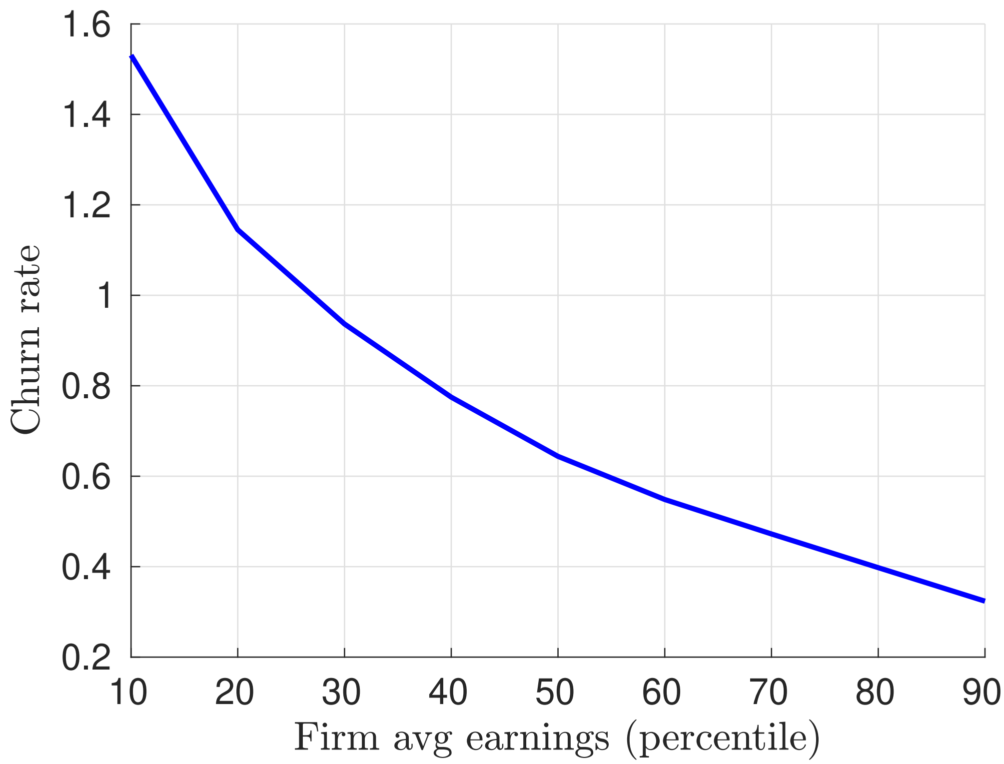

Figure 1: The mean of excess accession or "churn" rate (annual) at each pseudo-decile bin of the average firm-level wage distribution.

3.1 Worker-level estimation: Churn

While this strong negative relationship between churn rates and firm wages is suggestive of a churn ladder, this correlation can reflect differences in worker composition across firms without necessarily being related to individuals moving up a job ladder. To address this, we document that looking within continuous employment spells, workers move up a wage ladder and down a churn ladder, from low-wage and high-turnover firms into higher-wage and low-turnover firms with each job transition. By looking within a worker, this movement doesn't reflect sorting on the characteristics of the worker. Further, we look within a continuous employment spell, which controls for the initial starting point out of unemployment. Rather, the pattern we find on the change in earnings and firm turnover with each job transition reflects the worker's marginal gain in wage from searching on the job and making a transition, and the firm's implied tradeoff between match duration and its share of match surplus for the same worker.

We run the following regression on our sample of job transition observations:

$$X_{ij} = \beta_0 + \sum_{k=1}^{N_g} \beta_k \mathbb{I}_{j=job_k} + \gamma_g + \varepsilon \tag{3.1}$$

Where the X_{ij} is one of three measures, $X_{ij} = \bar{w}_{ij}$, $X_{ij} = \bar{L}_{ij}$, are the average earnings and number of employees at the firm j in the year before worker i separates from it. $X_{ij} = EAR_{ij}$ is the excess accession rate of firm j in the year before the worker i's separation. The index $\mathbb{I}_{job_k}$ is equal to one if the job is the kth job within a spell of employment, indexed by g. Including employment spell fixed effect γ means these coefficients are relative to the average for the employment spell. If a worker's first employer out of non-employment is quite good, e.g. low churn or high wage, then we want to look at that as the base from which we look for changes.

The coefficients β_k reflect the change in churn, wage or size from the transition to the k^{th} job in a spell from the baseline level of churn in the first job (which is normalized to 0). The inclusion of the employment spell fixed effect is particularly important because there is substantial heterogeneity not just across individuals, but across employment spells. By controlling for that heterogeneity, we isolate the effect of moving up the job ladder in a strict sense and quantify the change in the firm's characteristic with each move on the worker's job ladder. It is important to note that the employment spell fixed effect does not control for heterogeneity that creates differences in employment duration. That is, there is still a potential composition effect due to the survivorship of employment spells necessary to reach higher consecutive job transitions.

Figure 2 plots the coefficients from the above regression, giving us a visual representation of the churn ladder, size ladder and earnings ladder. Each begins at 0, which is a normalization because we are running this with fixed effects. Another way to see the normalization is the workers take their first job as a random draw on the distribution and movements thereafter on the ladder is based on changes from this position.

For graphical comparability, we look at logs of each variable, log earnings, log size and the log churn rate. The most striking trend is the monotonic decline in the log churn rate (the black dashed line), which falls by approximately 40 log points after seven job transitions. This downward trajectory indicates that with every move, workers successfully transition into firms characterized by higher stability and lower turnover. This descent of the churn ladder occurs simultaneously with a steady climb up the earnings ladder (the blue dotted line), confirming that workers generally move toward higher-paying employers over time. Here, the rise is about 30 log points over 7 transitions.

In contrast to the clear slopes of earnings and churn, firm size (the red dashed line) remains relatively flat across the job sequence. This suggests that while workers are effectively optimizing for higher pay and better retention, firm size itself is a less direct indicator of a job's position on the ladder compared to churn and wages. It rises, as would be predicted by theory, but only very modestly. If the central logic of the job ladder is that firms trade off higher earnings for better retention, then the black and blue lines are necessary outcomes. Size is a secondary implication and is also going to be affected by mitigating factors like differences in the production function hence its much weaker relationship.

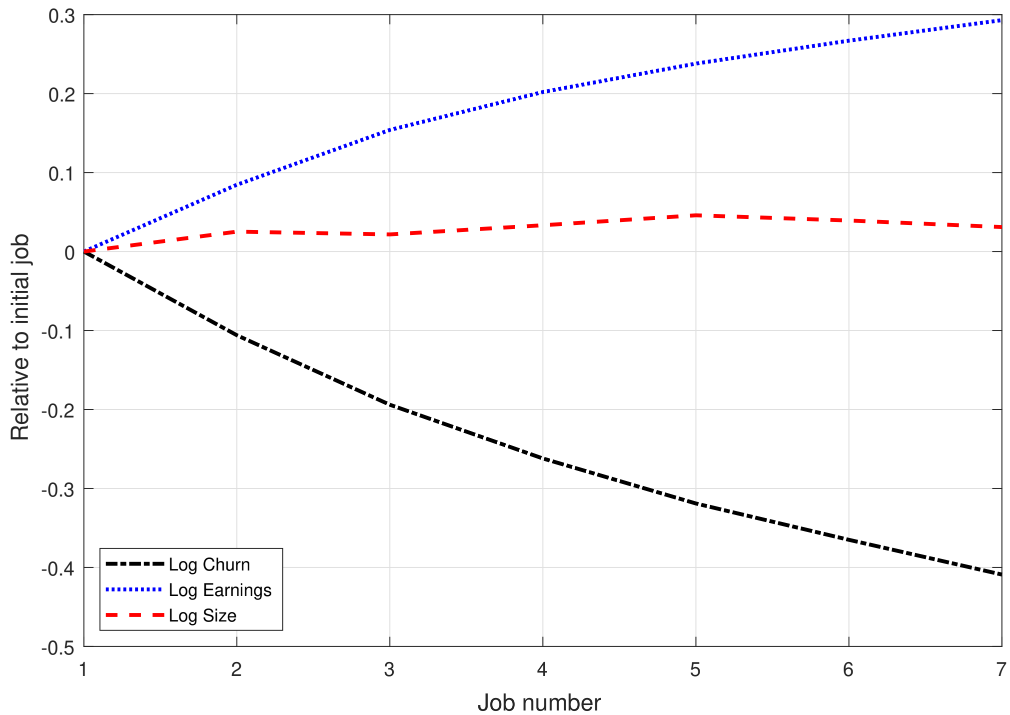

Figure 2: The firm's churn rate, size, and average earnings over the index of job transitions within a worker's employment spell. Each measure is normalized to 0 for the initial job in the spell.

3.2 Empirical tests for separation heterogeneity within wage

While the worker-level regressions in the previous section confirm the existence of a churn ladder, they do not tell us why it is so steep. In the canonical job ladder framework, the negative relationship between wages and churn has a single driver: poaching. High-wage firms lose fewer workers to competitors, and this differential poaching rate is the entire source of cross-firm variation in turnover. But the very steep gradient we observe–churn rates falling by nearly 40 log points over seven job transitions–could actually stem from several potential sources, broadly split into firm-level policies or the composition of workers the firm employs.

This distinction matters because it points to two fundamentally different sources of the churn ladder. One possibility is that the churn gradient is entirely a firm-side phenomenon: firms set wage and retention policies jointly, so that conditional on wage, all workers at a given firm face the same separation risk. Under this view, knowing a firm's wage rank tells you everything about its turnover, and the distribution of tenure within a wage bin should follow a simple exponential. The other possibility is that workers themselves differ in their underlying propensity to separate to nonemployment—some workers are inherently more attached to employment than others—and these types sort endogenously across the wage distribution as low-separation workers accumulate at high-wage firms through longer employment spells. This worker-side heterogeneity would generate excess dispersion in tenure within wage bins, well above what a constant separation rate would predict. The two tests we develop below use precisely these conditional moments of tenure—its mean and variance at each wage level—to distinguish between these explanations.

To set up both tests, first, imagine there is no worker-level heterogeneity in δ_j. Then, there should be no within-wage variation in separation probabilities. For notational simplicity, suppose the only reason for separation is into non-employment, $\delta(w)$.[8]

So in this case, without worker heterogeneity, the expected tenure at a given wage level should be $\tau(w) = \frac{1}{\delta(w)} = \frac{1}{\delta e^{-\epsilon w}}$. Now, suppose that the true data-generating process involves heterogeneous workers with two separation types, δ_1 and δ_2. Let $\mu_j(w) = \frac{h(w,\delta_j)}{g(w)}$, then the separation to non-employment can be written as

$$\begin{aligned}\delta(w) &= e^{-\epsilon w}\left[\mu(w)\delta_1 + (1-\mu(w))\delta_2\right] \\ &= \mu(w)\delta_1 e^{-\epsilon w} + (1-\mu(w))\delta_2 e^{-\epsilon w}\end{aligned}$$

This implies the expectation of tenure at a wage w,

$$\tau(w) = \mu(w)\frac{1}{\delta_1 e^{-\epsilon w}} + (1-\mu(w))\frac{1}{\delta_2 e^{-\epsilon w}} \ .$$

First-moment test: Now we can compare our null hypothesis, $\mu(w) = 1\ \forall w$, to a case

[8]Note, of course, that without on-the-job search there's no mechanism for dispersion in the wage offer distribution in our model anymore. So there would only be a single offered wage and a single variance of tenure exponentially distributed. For purposes of exposition, we are temporarily suspending our understanding of the Diamond Paradox.

where δ incorporates heterogeneity $\delta(w) = \mu(w)\delta_1 e^{-\epsilon w} + (1-\mu(w))\delta_2 e^{-\epsilon w}$. Since $\frac{1}{x}$ is a strictly convex function, by Jensen's inequality the expectation of tenure in the no-heterogeneity case is strictly smaller than the expectation of tenure in the case with heterogeneity. That is,

$$\frac{1}{\delta_0 e^{-\epsilon_0 w}} < \mu(w)\frac{1}{\delta_1 e^{-\epsilon w}} + (1-\mu(w))\frac{1}{\delta_2 e^{-\epsilon w}}.$$

Where δ_0, ϵ_0 are the average separation and firm effects in which all heterogeneity in separation is driven by firm heterogeneity and there is no variance within a firm wage level. Our test is then against the null hypothesis that the expectation of tenure is consistent with the no-heterogeneity case. We can reject the null if the empirical mean of $\frac{1}{\delta(w)}$ is greater than $\frac{1}{\delta e^{-\epsilon w}}$ from the model calibration at each w.

This first-moment test can be formulated as $H_0 : E[\tau|w] = \frac{1}{\hat{\delta}(w_i)}\ \forall w_i \in \{w_1, \ldots, w_N\}$. We test this on a finite set of w_i, where we have estimated the empirical $\hat{\delta}(w_i)$s.

At each w_i, $\frac{1}{\hat{\delta}_i} - \hat{\tau}_i$, where we use $\hat{\delta}_i, \hat{\tau}_i$ to mean the separation rate and tenure at each wage level w_i. Under the null hypothesis, the match tenure is driven by a homogeneous Poisson process, so it is distributed exponentially with rate parameter δ_i, $\tau \sim exp(\delta_i)$. This means $2\delta_i\tau \sim \chi^2_2$ and hence our test statistic at each w_i is

$$2\hat{\delta}_i\left(\hat{\tau}_i - \frac{1}{\hat{\delta}_i}\right) \sim \chi^2_{2n_i-1}\ .$$

To make this into a joint hypothesis, notice that $\delta(w_i)$ is independent across each w_i. Thus, the joint hypothesis test is distributed $\chi^2_{2\sum_{i=1}^N n_i - N}$

$$\sum_{i=1}^{N} 2\hat{\delta}_i\left(\hat{\tau}_i - \frac{1}{\hat{\delta}_i}\right) \sim \chi^2_{2\sum_{i=1}^N n_i - N} \tag{3.2}$$

This problem gets more complicated in our true model with on-the-job search, in which the separation rate $s(w)$ from a firm of wage w is

$$\begin{aligned} s(w) &= \delta(w) + (1 - F(w))\lambda_e + \lambda_r \\ &= e^{-\epsilon w}\left[\mu(w)\delta_1 + (1 - \mu(w))\delta_2\right] + (1 - F(w))\lambda_e + \lambda_r \\ &= \mu(w)\delta_1 e^{-\epsilon w} + (1 - \mu(w))\delta_2 e^{-\epsilon w} + (1 - F(w))\lambda_e + \lambda_r \ . \end{aligned}$$

This is a competing risks problem such that the time before separation is distributed as an exponential with rate $\delta(w) + (1 - F(w))\lambda_e + \lambda_r$.

However, the endogenous function, $F(w)$, moving freely, presents two problems. First, the null hypothesis for each w_i becomes difficult to characterize, as a function of w_i itself. More difficult, however, is that each w_i is no longer independent. But, we can simplify this by invoking our empirical result that separations to non-employment are a constant fraction of all separations, which we will call $1 - \ell$. Under the null hypothesis that firm heterogeneity is the data generating process, $s(w) = \frac{\delta(w)}{1-\ell}$ and thus our null for the mean of tenure is proportional to our earlier H_0.

Second-moment test: Similarly, we can devise a statistical test using the variance of tenure and not just the mean. This has a nice intuition that worker heterogeneity in separation rates tends to be largest at the bottom of the wage distribution, pushing up dispersion in tenure. However, these also have the highest separation rates, pushing down the dispersion of tenure. Hence, if there is no heterogeneity, only the latter force operates.

To formalize this logic, again suppose that separations to non-employment are the only determinant of tenure, i.e. no on-the-job search. Then, the variance of tenure as a function of wage is

$$\begin{aligned} \mathrm{Var}(\tau(w)) &= \mu(w)^2 \frac{1}{(\delta_1 e^{-\epsilon w})^2} + (1 - \mu(w))^2 \frac{1}{(\delta_2 e^{-\epsilon w})^2} \\ &= \mu(w)^2 (\delta_1 e^{-\epsilon w})^{-2} + (1 - \mu(w))^2 (\delta_2 e^{-\epsilon w})^{-2} \end{aligned}$$

If $\mu(w) = 1$ so that there is no heterogeneity, then the variance of tenure follows as a simple exponential distribution at each w. Hence we can test against that null hypothesis that $\text{Var}(\tau(w)) = \frac{1}{(\delta_1 e^{-\epsilon w})^2}$ $\forall w$. Of course, we can actually observe $\delta(w)$ in the data, so our null hypothesis is

$$H_0 : \text{Var}(\tau|w_i) = \frac{1}{\hat{\delta}_i^2} \ \forall w_i \tag{3.3}$$

We can again test on a finite set of $w_i \in \{w_1, \ldots, w_N\}$ and the empirical $\hat{\delta}_i$ estimated at each w_i. This yields a test statistic at each w_i that is the estimated variance at each i and the one constructed from separation rates at i using the estimated exponential distribution: $\widehat{var(\tau|w_i)} - \frac{1}{\hat{\delta}_i^2}$.

Incorporating on-the-job search, the variance of tenures is again somewhat more complicated:

$$\text{var}(\tau|w) = \frac{1}{(\delta(w) + (1 - F(w))\lambda_e + \lambda_r)^2} = \frac{1}{(s(w))^2}$$

Incorporating our empirical finding that $s(w) = \frac{\delta(w)}{1-\ell}$, our null hypothesis is now that $\text{var}(\tau|w) = \frac{(1-\ell)^2}{\delta(w_i)^2}$ and the rest of the analysis is the same. The test statistic,

$$\widehat{\text{var}(\tau|w)} - \frac{(1-\ell)^2}{\hat{\delta}_i^2}$$

The distribution of this test statistic are difficult to characterize because it is the product of two $\chi^2_{2n_1}$ distributed variables. The joint hypothesis is then distributed as the sum of such products. Hence, we bootstrap the distribution and use this to generate p-values.

In Figure 3, we present the test statistics for both second and first moment tests across the wage domain. The joint test statistic of each has a p-value indistinguishable from 0 both because our sample size is so large but also because the deviation at each point is so large. In the first moment, the actual tenure is considerably longer than that implied by a constant separation rate within firms. For the second moment, there is considerably more dispersion in tenure than would be predicted. Both are suggestive of workers with different separation rates within each wage bin.

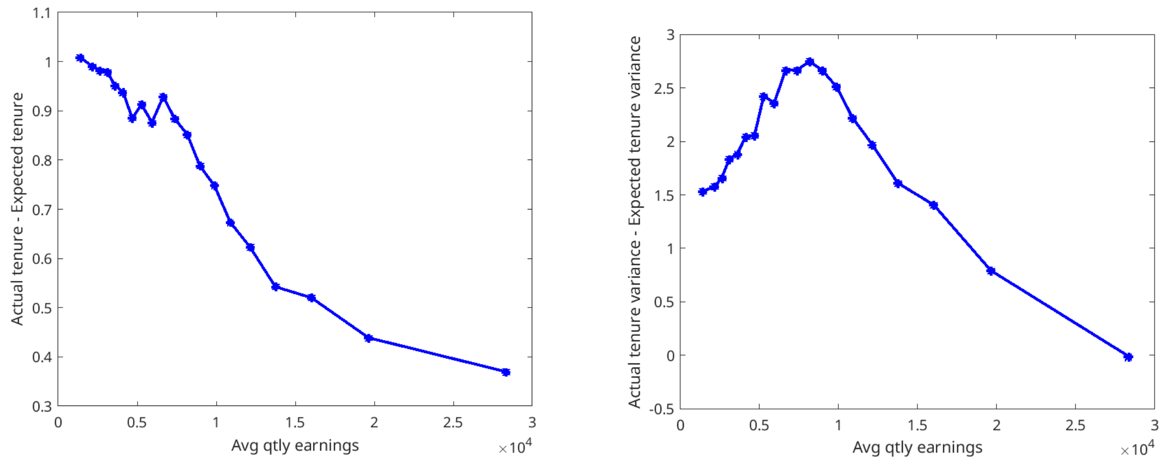

Figure 3: First moment (LEFT) and second moment (RIGHT) tests for the presence of worker heterogeneity in separation rates. By both tests we reject uniformity.

Together, these rejections establish that firm-level wage policy alone cannot account for the steepness of the churn ladder. The data require a model in which workers bring their own separation risk to the firm—and in which that risk sorts endogenously across the wage distribution as low-separation workers survive longer employment spells and climb higher on the ladder. This is not merely a statistical footnote: the magnitude of the deviations in Figure 3 is large, not just statistically significant, which means the quantitative contribution of worker heterogeneity to the churn gradient is likely substantial. In what follows, we build a model that incorporates exactly this structure—heterogeneity in both worker and firm separation rates alongside the standard poaching mechanism—and use it to decompose how much of the observed churn ladder each source explains.

4 How much churn in a job ladder

We start with a framework based on Burdett and Mortensen (1998). This job ladder model has a simple flow-balance condition that characterizes the distribution of firm size via hires and separations so that for a firm offering wage w to the worker, its size L is characterized by:

$$(\delta + \lambda_e(1 - F(w))L(w) = \lambda_u u + \lambda_e G(w)(1 - u) \tag{4.1}$$

where δ is the separation rate, λ_e is the rate of contact of employed workers, F is the distribution of offers from other firms, G the distribution of values among currently matched workers. We take as u the mass of unemployed workers, λ_u is their rate of contacting firms, and the overall measure of workers and firms as 1. The left-hand side is the number of workers exiting from a match with wage w and the right-hand side is the number of workers entering it. In this simple version, wage w summarizes the match value. In equilibrium the size of the firm, $L(w)$, is characterized by the firm indifference condition $L(w)(p - w) = \bar{\pi}$ where $\bar{\pi}$ is some constant level of profit and p is the productivity level of all firms. While we will not impose this equilibrium condition, it provides some useful intuition as to why the job ladder's wage-churn relationship is so general. Here, wages influence worker retention and thereby affect firm size and firm churn rates. A firm with lower turnover rates maintains a larger firm size by keeping its workers longer, resulting in a higher expected surplus from the same per-worker profit flow, $p - w$. Although higher-wage firms earn less flow profit on each worker, they are compensated by reduced turnover.

Given the implied assumption that workers transit to higher-paying firms, $1 - F(w)$ defines the dominating offers, and known distributions and parameters, Equation 4.1 defines a relationship between wages and flows. This implies that one can take the empirical distribution of size and excess hires and infer the distribution of offered values or vice-versa.

Specifically, we can define the per-worker excess hires rate, $c(w)$, the hires above and beyond the growth rate of the firm, as

$$c(w) = \delta + \lambda_e(1 - F(w)) = \frac{\lambda_u u + \lambda_e G(w)(1 - u)}{L(w)} \ . \tag{4.2}$$

Because we are considering firms in steady-state, every hire is an "excess" hire and all gross worker flows are churn. Thus, $c(w)$ equals both the hire rate and the separation rate, whereas more generally it would be the minimum of the two. To obtain an empirical relationship from this model, the first equation, $c(w) = \delta + \lambda_e(1 - F(w))$, is most useful.

Specifically, it is easy to see the monotonicity of churn on wages. Taking the derivative of $c(w)$ yields

$$c'(w) = -\lambda_e f(w)$$

which is weakly negative everywhere because $f(w)$, being a density, is non-negative everywhere. Both empirically and in one of the most famous results of Burdett and Mortensen (1998), there is positive mass on $f(w)$ over the entire domain of wages. This means that $c(w)$ strictly declines with wages.

Furthermore, we can define the flows in and out of the distribution below $F(w)$, given an earnings distribution $G(w)$. This flows equation relates the distributions $F(w)$ and $G(w)$ by:

$$F(w) = \frac{G(w)(\delta + \lambda_e)(1 - u)}{\lambda_u u + \lambda_e(1 - u)G(w)} \ . \tag{4.3}$$

This holds by equating the mass of workers falling down the job ladder to those who are climbing the job ladder, the same logic as presented in Jolivet et al. (2006). This last relationship is particularly useful empirically because $G(w)$ can be directly observed in the data, as can δ.

4.1 Heterogeneous separations and the job ladder

Our empirical results suggest that separations are a key component of churn, so we allow δ to vary across both workers and firms. Focusing first on worker heterogeneity, the intuition is that different workers fall off the job ladder into nonemployment at different rates.Thinking first about the former, the intuition is that different workers will fall off the job ladder by moving to nonemployment at different rates. This implies a pattern of sorting in which higher-wage firms at the top of the ladder accumulate low separation rate types. Low-wage firms that hire more directly from unemployment have more separations as the unemployed consist of many high-separation types. Going up the wage ladder, firms consist of more low-separation types who were able to remain employed for consecutive job-to-job transitions as high-separation types move back to unemployment.

The intuition for firm-level differences in δ is that firms may bundle the wage and the stability of the job they offer. There are several reasons this may happen, for instance, if there are diminishing returns from monetary remuneration, longer-lived jobs might be more valuable. Mechanically, then we can build in a relationship whereby high-wage employers also offer lower δ values.

In addition to these two dimensions of heterogeneity, to empirically match the distribution of worker turnover and the relationship between wages and churn in job transitions, we must account for the possibility of workers moving to jobs that are dominated by the previous employer's wage level. We introduce a reallocation shock with an arrival rate of λ_r, in which a worker moves to a new job drawn from the entire offer distribution F regardless of the wage[9].

To add these three concepts, suppose the distribution of workers' separation rates is $D(\delta)$ and each worker gets a draw from this distribution, $\delta_i \in [0, \bar{\delta}]$, that is fixed forever. The firm side heterogeneity is defined by an elasticity ϵ_δ, such that it scales the separation rate by $e^{\varepsilon_\delta w}$. This implies that conditional on δ_i, one could derive the same job-churn relationship from Equation 4.2 but at the firm level, however, a particular wage firm will have a $\delta(w)$ that depends on its accumulation of δ_i types and its own separation policy defined by ε_δ. The separation rate at wage w is therefore $\delta(w) = e^{\varepsilon_\delta w} \int_0^{\bar{\delta}} s dD(s|w)$.

The reservation wage of every worker at a firm paying w is still w, although there is the possibility of a reallocation shock with probability λ_r. This means the mass of potential poaching offers is still $(1 - F(w))$ for on-the-job search. This comes because we forced all firm-side heterogeneity to be perfectly correlated with the wage, so in moving up in wages, the firm component of the separation risk also gets more favorable.

[9] This reallocation or "Godfather" shock is an offer that the worker cannot refuse and induces a move to anywhere in the offer distribution regardless of the worker's current wage.

This means that our new expression for churn, $c(w)$, is now

$$c(w) = e^{\varepsilon_\delta w} \int_0^{\bar{\delta}} s dD(s|w) + \lambda_r + \lambda_e(1 - F(w)) \tag{4.4}$$

We now present a modified version of Equation 4.3, including both heterogeneity in workers and firms and our new reallocation shock, λ_r.

$$\int_{\underline{w}}^{w} \int_0^{\bar{\delta}} (se^{\varepsilon_\delta z} + \lambda_r(1 - F(w)) + \lambda_e(1 - F(w)))(1 - u)h(z, s)dsdz =$$
$$\lambda_u uF(w) + F(w) \int_w^{\bar{w}} \int_0^{\bar{\delta}} \lambda_r(1 - u)h(z, s)dsdz \tag{4.5}$$

To see how this compares to our initial setup, just set $\delta_i = \delta, \varepsilon_\delta = 0$ and we recover G as the marginal distribution, integrating over δ_i and from $z = 0$ to w: $G(w) = \int_{\underline{w}}^{w} \int_0^{\bar{\delta}} h(z, s)dsdz$.[10] Note that we are integrating over wages, using z to index them. Hence, the joint relationship between δ and w appears on the left-hand side of Equation 4.5 as the accumulation of firms' separation policy below wage w.

4.2 Mapping model to data

Suppose wages are defined on a discrete interval, $\{w_1, \ldots, w_W\}$ and there are a discrete set of types and $\{\delta_1, \ldots, \delta_\Delta\}$, which defines discrete densities h.[11] Then for wage value w_k we

[10] Using this style of notation, we did not need to use the joint density h, instead we could use conditional distributions, $G(w|\delta)$ and we could have written equation 4.5 as

$$\int_0^{\bar{\delta}} (se^{\varepsilon_\delta w} + \lambda_r(1 - F(w)) + \lambda_e(1 - F(w)))(1 - uG(w|s)ds = \lambda_u u(s)F(w) + F(w) \int_0^{\bar{\delta}} \lambda_r(1 - u))G(w|s)ds$$

[11] We will abuse notation somewhat, continuing to use notation F, h to describe the distributions and densities associated with w although they are now discrete instead of continuous above.

have the following:

$$\sum_{i=1}^{k}\sum_{j=1}^{\Delta}(\delta_j e^{\varepsilon\delta w_i} + (\lambda_r + \lambda_e)(1 - F(w_k)))(1-u)h(w_i,\delta_j) = \lambda_u \sum_{j=1}^{\Delta} u_j p_j F(w_k) + F(w_k) \sum_{i=k+1}^{W}\sum_{j=1}^{\Delta}\lambda_r(1-u)h(w_i,\delta_j) \tag{4.6}$$

Where the wage-specific separation rate is $\delta(w_k)$. That is, the effective separation rate for a firm at wage w_k is the worker-weighted average of separation rates δ_j for each j type of worker at wage w_k.

We can use Equation 4.6 to solve for $F(w_k)$

$$F(w_k) = \frac{\sum_{i=1}^{k}\sum_{j=1}^{\Delta}(\delta_j e^{\varepsilon\delta w_i} + \lambda_r + \lambda_e)(1-u)h(w_i,\delta_j)}{(\lambda_r+\lambda_e)\sum_{i=1}^{k}\sum_{j=1}^{\Delta}(1-u)h(w_i,\delta_j) + \lambda_u\sum_{j=1}^{\Delta}u_j p_j + \sum_{i=k+1}^{W}\sum_{j=1}^{\Delta}\lambda_r(1-u)h(w_i,\delta_j)} \tag{4.7}$$

And then our densities $h(w_i, \delta_j)$, $W \times \Delta$ values, are defined as a system of $W \times 2$ equations:

$$\delta(w_i) = \left(\sum_{j=1}^{\Delta}\delta_j \frac{h(w_i,\delta_j)}{g(w_i)}\right) e^{\varepsilon\delta w_i} \tag{4.8}$$

$$g(w_i) = \sum_{j=1}^{\Delta} h(w_i,\delta_j) \tag{4.9}$$

$$\text{for } i \in \{1, ..., W\}$$

and we can solve for the probability of each type by using the steady-state unemployment rate:

$$u = \sum_{j=1}^{\Delta} u_j p_j$$

Where u_j is the steady state for type j workers, which involves integrating over employment. Define the conditional density $g(w_i|\delta_j) = \frac{h(w_i,\delta_j)}{\sum_{s=1}^{W} h(w_s,\delta_j)}$. Then for type j, their unemployment rate is

$$u_j = \frac{\delta_j \sum_{i=1}^{W} e^{\varepsilon_\delta w_i} g(w_i|\delta_j)}{\lambda_u + \delta_j \sum_{i=1}^{W} e^{\varepsilon_\delta w_i} g(w_i|\delta_j)} \tag{4.10}$$

Of course, with $\Delta = 2$ types, we can use Equations 4.8, 4.9 and 4.10 to solve for this entire system, as given in Appendix C.

Using Bayes' rule:

$$\Pr[w = w_i|\delta = \delta_j] = \Pr[\delta = \delta_j|w = w_i]\frac{\Pr[w = w_i]}{\Pr[\delta = \delta_j]} \tag{4.11}$$

$$\frac{h(w_i, \delta_j)}{\sum_s h(w_s, \delta_j)} = \frac{h(w_i, \delta_j)}{\sum_j h(w_i, \delta_k)} \frac{g(w_i)}{(1 - u_j)} \tag{4.12}$$

and then

$$\sum_s h(w_s, \delta_j) = \sum_j h(w_i, \delta_k) \frac{(1 - u_j)}{g(w_i)} \tag{4.13}$$

5 A Calibrated job ladder model

In this section, we calibrate our model and compare the model-implied relationship between churn and wages with the corresponding patterns in the data. Our baseline is the full model, which combines on-the-job search, reallocation shocks, wage-dependent separations, and permanent heterogeneity in workers' separation rates. We show that this richer specification can jointly account for the worker churn ladder, the firm-level churn gradient, the distribution of job tenure, and the composition of separations into job-to-job and nonemployment transitions. We then use two counterfactuals to isolate the role of the different mechanisms. The first is the Burdett-Mortensen benchmark, which shuts down wage-dependent separations for a constant separation rate. The second removes worker heterogeneity while retaining wage-dependent firm separation policies and reallocation. Together, these counterfactuals show what pure on-the-job search can generate, and why worker heterogeneity is necessary to match the joint distribution of churn, tenure, and the firm distribution in the data.

We set the model to quarterly frequency to match the data. We take the normalized empirical CDF of firm wages as a direct input for the employment wage distribution, $G(w)$. In the full model, we then solve the system in Equations 4.7-4.10 to recover the offer distribution $F(w)$, the joint distribution of wages and worker separation types in employment, the worker separation parameters δ_j, and the reallocation rate λ_r. Holding fixed the population shares of the two worker types, p_1 and p_2, the remaining parameters are chosen to minimize the distance between the model and the data along several dimensions.

Table 1: Calibrated Model parameters

Parameter	Value	Target
p_1	0.8	Ahn et al. (2023)
p_2	0.2	
λ_u	0.084	Quarterly NE rate 0.084
λ_e	0.045	est. - worker ladder
λ_r	0.023	est. - worker ladder
$\varepsilon_\delta w$	-0.195	est. - worker ladder
s_1	0.062	est. - tenure distribution
s_2	0.73	

We choose the five estimated parameters to minimize the distance between model and data for the following moments: (i) the worker churn ladder, measured as the change in churn experienced at each job index; (ii) the share of workers with less than one year of tenure across the firm wage distribution; (iii) the fraction of separations that are to nonemployment rather than to other employers across the firm wage distribution; (iv) the overall churn rate across the firm wage distribution; and (v) the average separation rate of all workers.

5.1 Baseline results: the full model

We begin with the full calibrated model. Relative to the standard job ladder, this model allows separations to nonemployment to vary with wages through two channels: a firm-level wage-dependent separation policy and permanent worker heterogeneity in ex ante separation risk. These two forces jointly lower the effective separation rate at high-wage firms and raise it at low-wage firms. Low-separation workers survive longer employment spells, accumulate more opportunities to move up the ladder, and therefore sort toward high-wage firms. High-separation workers are disproportionately concentrated at low-wage firms and in unemployment. This endogenous sorting amplifies the decline in churn over the wage distribution, even though search remains random conditional on contact.

Underlying the endogenous relationships in our model, we backed out $F(w)$ via a system of Equations 4.7-4.10. As expected, $G(w)$ first order stochastic dominates $F(w)$, which must happen in any job ladder model. The sorting across types and wage-specific firm policies make this a bit more extreme: to get mass at the top of the earnings distribution requires relatively few offers because these jobs are mostly filled by low separation type workers and there are very few outflows to nonemployment or other jobs. Hence, $F(w)$ approaches 1 quite quickly.

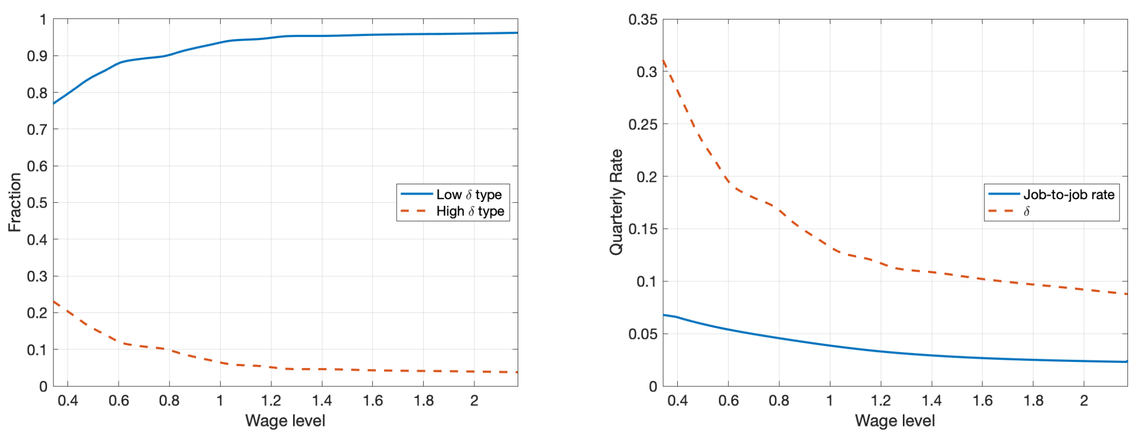

Figure 4: LEFT: The fraction of low- and high-separation-rate worker types at each level of the firm wage distribution in the calibrated model. RIGHT: The quarterly rate of job-to-job transitions (including both on-the-job search and reallocation) and the average separation rate $\delta(w)$ integrated over all worker types at each level of the firm wage distribution in the calibrated model.

To see the endogenous sorting, the left panel of Figure 4 plots the composition of worker types and separation probabilities over the firms' average wage distribution. At the lowest wage firms, their workforce is mostly composed of high separation rate workers: they are small and mostly recruit from unemployment, which low δ-workers rarely experience. At the top of the wage distribution are mostly low δ-workers, but who are joined to some extent by some lucky high δ types.

The result of this sorting, along with the firms' elasticity ε_δ creates our differences in $\delta(w)$ shown in the right panel of Figure 4 with the endogenous rate of job-to-job transitions. The job-to-job rate reflects the poaching hierarchy, and is quite steep in large part because of the steepness of $F(w)$. There are a lot of job transitions out of very low wage firms as there is a lot of churn and a high density of offers. $\delta(w)$ is higher in level but remains proportionally constant with job-to-job transition rates, calibrated to our empirical findings on the share of separations. Its decline over w reflects the sorting of lower δ types to higher wage levels as well as the firm-level separation policy declining in wage.

Figure 5 shows two of our calibration targets. On the left panel, it compares the slope of the worker churn ladder in our model relative to our estimates in the data. The decline in churn with each job transition is similar in magnitude in our model, which is basically a function of the shape of the distribution of churn in our model. Workers mostly move up the job ladder, and these steps up coincide to the proper amount of change in churn due to the exogenous $\delta(w)$ and the endogenous job-to-job mechanism.

In the right panel, we show the distribution of short-tenured matches. As introduced in our empirical tests, while we might match the distribution of turnover with effects that vary purely with w, if all of the separation rates are homogeneous conditional on w, we would miss tenure. Hence, our match with Figure 5 shows that we have properly disciplined the worker heterogeneity.

At the worker level, the composition of workers who remain employed through several job transitions consists mostly of those with a low separation probability. This variation in composition mirrors the type of selection present in our regressions with employment-spell fixed effects. This sorting through the longevity of employment is present despite the fixed-effects specification controlling for heterogeneity across jobs within the employment spell and across workers.

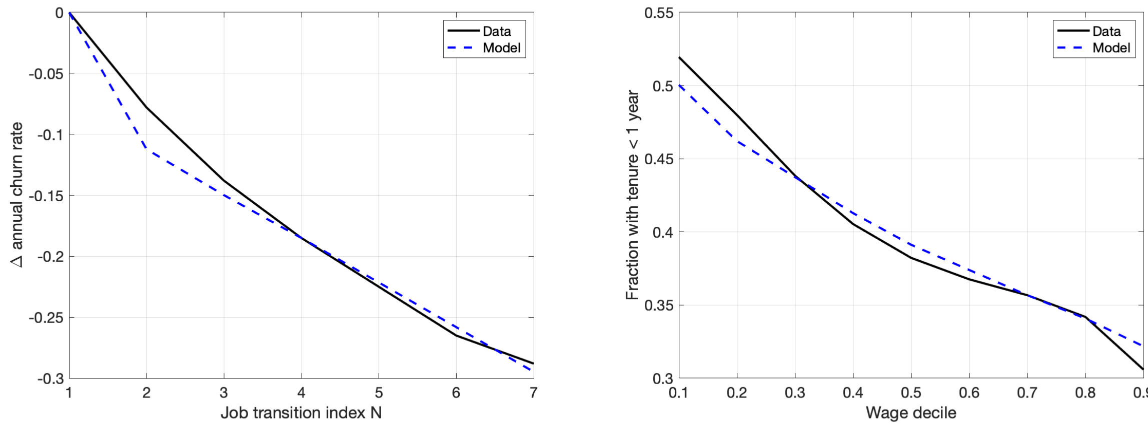

Figure 5: LEFT: The change in annual churn rate of the employer relative to the initial job across job ladder transitions in the data and the calibrated model. The average churn rate of the initial job is normalized to 0. RIGHT: The fraction of workers with less than one year of tenure at each decile of the firm wage distribution in the data and the calibrated model.

Figure 6 (LEFT) plots the annual churn rate against the earnings decile of the firm, comparing the data (solid black line) to the calibrated model with worker heterogeneity (dashed blue line). This is also a calibration target, and we match it well. Both lines begin around 1.2 at the lowest wage decile and decline monotonically to roughly 0.4–0.5 at the top of the distribution. The most notable feature of this figure, especially in contrast to Figure 4, is how closely the model now tracks the data across the entire wage distribution. The fit covers both of the extremes—the model replicates both the high churn rates at the bottom of the ladder and the much lower rates at the top of the firm average wage distribution. This stands in sharp contrast to our counterfactual exercises without heterogeneity and without wage-dependent separations. The tight fit reflects the two mechanisms we have added that vary separations over the wage distribution: worker composition shifts toward low-separation types as one moves up the wage distribution, and firms' own separation policies also become more favorable at higher wages. Together, these compress the churn rate at the top of the distribution.

In Figure 6 (RIGHT), the model also targets the fraction of separations into non-employment. Interestingly, this share is nearly constant over the wage distribution. While separations decline dramatically over the wage distribution, their composition in terms of

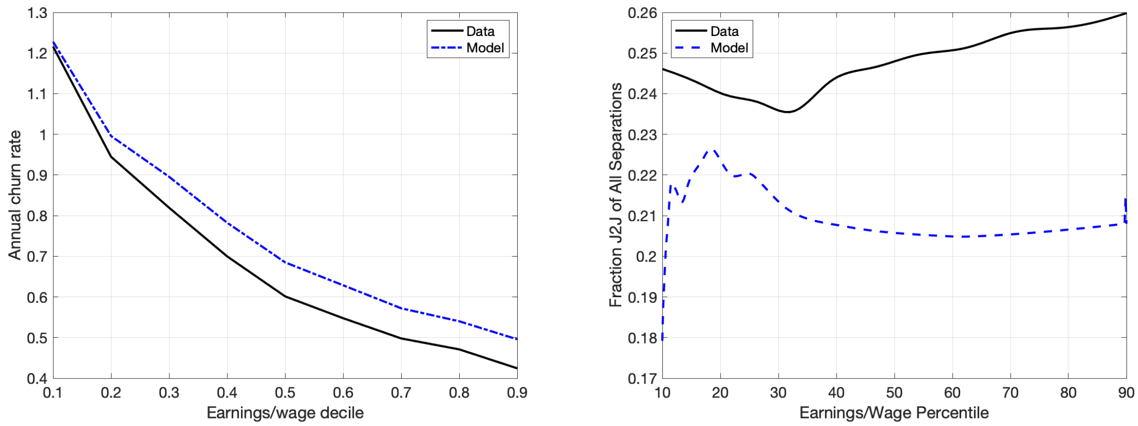

Figure 6: LEFT: The mean annual churn rate at each decile of the firm earnings distribution in the data and the calibrated model with worker heterogeneity. RIGHT: The fraction of separations to other employers (job-to-job) out of all separations at each percentile of the firm earnings distribution in the data and the calibrated model with worker heterogeneity.

the share leaving to other firms vs nonemployment stays roughly the same. This relationship between separations to nonemployment vs. job-to-job is key to identifying the magnitudes of firms' elasticity ε_δ and the magnitude of the selection effect for workers. If separations to non-employment are primarily generated by the firm elasticity instead of worker composition, then job-to-job transitions contribute too much to turnover at the top of the wage distribution. Figure 6 shows how the calibrated model matches this fraction quite well.

In the model, we can split the separation rate into its components: those that are due to the relative composition of worker types, the firms' elasticity ε_δ, and job-to-job transitions. Figure 7 shows this breakdown across the wage distribution. We see that each source contributes to the overall separation rate. Throughout much of the wage distribution, worker composition and firms' δ policy contribute roughly equally in terms of the decline in separations over the wage distribution. Each contributes about half of the decline in separation rate over wages. Relative to job-to-job transitions, separations account for most of the total separation rate both in level and in their slope over the wage distribution.

To summarize the decomposition presented in Figure 9, the total decline in churn from the lowest to the highest wage level is about 25 percentage points. What is striking about this decomposition is that the relative contributions of each source are not constant across the wage distribution—they shift meaningfully as one moves up the ladder. At the bottom of the wage distribution, worker composition accounts for roughly 60% of the separation rate, with firm δ policy contributing about one-third and job-to-job transitions making up the remainder. Moving toward the top of the distribution, the picture changes considerably: worker composition becomes the near-exclusive driver, accounting for about 75% of separations, while the firm δ policy contribution fades to essentially nothing. Job-to-job transitions account for the remaining quarter of separations at the top, reflecting the fact that even high-wage firms lose some workers to λ_r. Taken together, this shifting decomposition tells a coherent story: firm-level retention policies matter most for workers near the bottom of the ladder, where wages still vary enough to affect separation incentives, but the concentration of low-separation-rate workers at the top is so dominant that it crowds out the other mechanisms almost entirely.

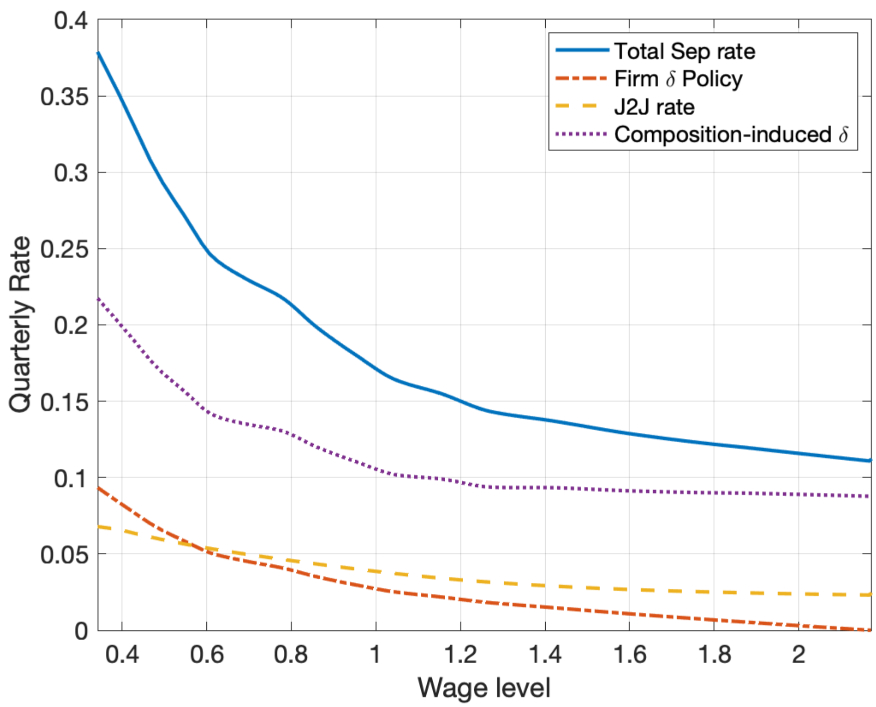

Figure 7: The contribution of worker types, firms, and job-to-job transitions to total turnover over the wage distribution.

Having shown that the baseline model matches the joint patterns in the data, we now use two restricted counterfactuals to isolate the contribution of its main ingredients. We first consider the canonical Burdett-Mortensen benchmark, which allows churn to vary with wages only through on-the-job search. We then study an intermediate counterfactual that retains wage-dependent firm separation policies and reallocation, but removes worker heterogeneity. These two exercises clarify why the full model is needed.

5.2 Counterfactual: The Burdett-Mortensen job ladder

for our first counterfactual scenario we consider the Burdett-Mortensen (BM) benchmark model, which sets the separation rate to a constant over the wage distribution. In this case, churn varies with wages only through poaching and job transitions, with a common separation rate to nonemployment for all workers. This benchmark is useful because it isolates how much of the observed churn-wage relationship can be generated by on-the-job search alone.

Figure 8 (LEFT) plots the worker churn ladder and on the (RIGHT) panel plots the average churn rate of firms in each decile of firm average earnings in the data, our calibrated baseline model, and the BM counterfactual. The firm-level churn distribution is where the BM benchmark fails most clearly. While the model with just job transitions can generate a negative relationship between earnings and churn, the elasticity is much higher in the data. While workers do move from higher churn to lower churn jobs as they move up the ladder, in a proportional sense, the overall magnitude of churn in the economy is too low.

To understand the mechanics of the simple model and why it fails to generate sufficient variation in churn, it is helpful to look at equation (4.2): $c(w) = \delta + \lambda_e(1 - F(w))$. It is clear that churn is increasing in the contact rate λ_e of employed workers searching on the job, but this increase depends on the offer distribution. As λ_e increases, the initial churn rate of employers hiring low-wage workers at the bottom of the job ladder increases, reflecting a high rate of poaching. But at the top of the job ladder, the churn rate of firms doesn't move much with λ_e since $(1 - F(w))$, the probability of the offer being higher than the

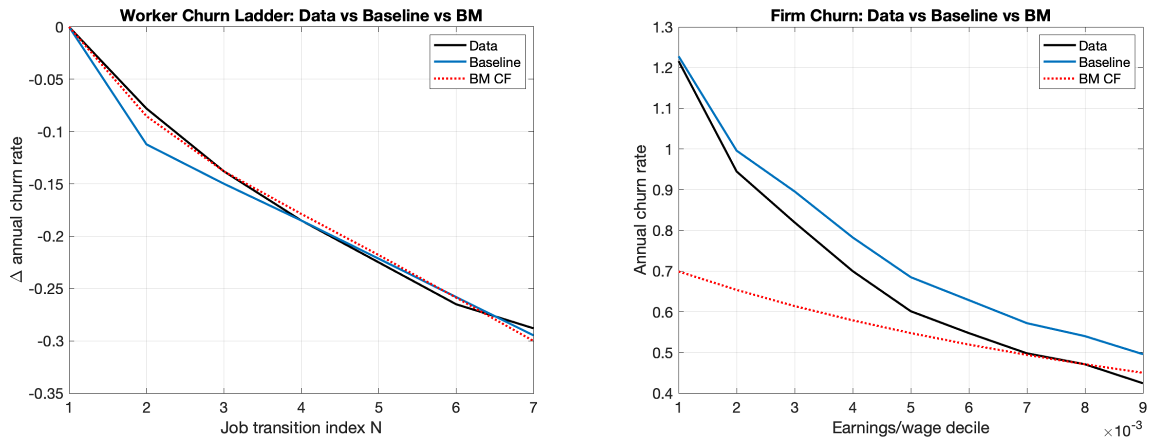

Figure 8: LEFT: The change in annual churn rate of the employer relative to the initial job across job ladder transitions in the data, the calibrated baseline model, and the Burdett-Mortensen counterfactual. The average churn rate of the initial job is normalized to 0. RIGHT: The mean annual churn rate at each decile of the firm earnings distribution in the data, the calibrated baseline model, and the BM counterfactual.

worker's current wage, is small. The reason for this high baseline value of churn is evident in the appearance of a constant δ in equation (4.2). Separation rates to nonemployment are another determinant of churn and in our baseline model, this rate is identical for all workers. As long as δ is a constant, it imposes this floor on churn even at the highest wage rates.

To achieve enough variation in churn across the job ladder and the firm distribution, separation rates to nonemployment must also be correlated with wages. Indeed, when we decompose worker separation rates at the firm into separations to other employers (EE) vs separations to nonemployment (EN), we found that EN rates are also negatively correlated with wage rank and positively correlated with churn.

The distribution of job tenure in the Burdett-Mortensen model shows how constant separation rates limit the fit of the model. Figure 9 plots the fraction of workers with tenure under one year across firm wage deciles. The BM counterfactual predicts too many short-tenure matches over most of the wage distribution, with the discrepancy pronounced at higher-wage firms. This failure follows from the model's central restriction: without reallocation shocks or heterogeneity in separation risk, the only force generating churn differences across firms is poaching along the job ladder. As a result, the model can produce declining churn with wages, but only by making high-wage firms increasingly insulated from separations. That mechanism is too weak to generate the observed dispersion in tenure while also matching the firm-level churn gradient.

The composition of separations provides even sharper evidence against the BM benchmark. The right panel of Figure 9 plots the fraction of all separations that are job-to-job across the wage distribution. In the data, the J2J share of separations is roughly flat. The BM counterfactual predicts the opposite: a steep decline, from above the data at the bottom to essentially zero at the top. This reversal is the direct implication of a model in which separations differ across firms only because of poaching. At low-wage firms, nearly all separations are attributed to workers moving up the ladder, while at high-wage firms the absence of better outside offers drives the J2J margin toward zero. But that is not what the data show. High-wage firms see proportional declines in separations to nonemployment, in lock step with the drop in job-to-job separations. Similarly, low-wage firms' elevated churn is not primarily a poaching phenomenon. The BM model therefore gets the level and slope of firm churn wrong for the same reason that it distorts the EN/EE decomposition: it lacks the nonemployment and heterogeneity margins that, in the baseline, generate substantial turnover at low-wage firms without forcing the entire adjustment onto job transitions.

5.3 The role of worker heterogeneity

Our model generates $\delta(w)$ through two distinct mechanisms. The first is a firm-level policy, parameterized by the elasticity ε_δ, that ties the separation rate directly to the wage offer. The second is latent worker heterogeneity in ex ante δ, which sorts low-separation workers toward high-wage firms through the natural selection of longer employment spells and repeated on-the-job transitions. Section 3.2 provided the empirical basis for both channels: we rejected the null that worker heterogeneity plays no role in separation rate dispersion conditional on wage.

To quantify the contribution of worker heterogeneity, we simulate a counterfactual that removes it entirely. We collapse both types to the population-average separation rate and

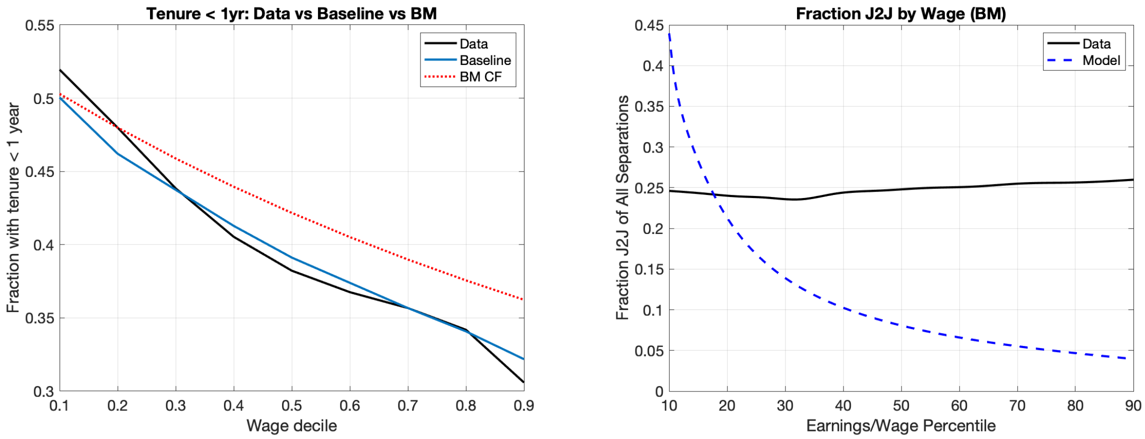

Figure 9: LEFT: The fraction of workers with less than one year of tenure at each decile of the firm wage distribution in the data, the calibrated baseline model, and the Burdett-Mortensen counterfactual. The BM counterfactual produces too many short-tenure matches throughout the upper part of the wage distribution. RIGHT: The fraction of all separations that are job-to-job transitions at each decile of the firm earnings distribution in the data and the BM counterfactual. The BM counterfactual produces too many separations to nonemployment relative to job transitions at the top of the wage distribution.

re-calibrate, relying solely on the firm-level elasticity ε_δ to replicate the $\delta(w)$ gradient. This restricted version resembles slippery job ladder specifications such as Pinheiro and Visschers (2015) and Jarosch (2023), which generate churn dispersion entirely through firm-level policies. While parsimonious and theoretically coherent, it cannot simultaneously match the churn ladder and the cross-sectional distribution of tenure — a tension the heterogeneous-worker baseline resolves through the additional degree of freedom that endogenous compositional sorting provides.

Figure 10 (LEFT) compares the worker churn ladder across the data, the baseline model, and the no-heterogeneity counterfactual. The no-het counterfactual tracks the slope of the ladder reasonably well: the rate of decline in churn across job transitions is broadly similar across all three, with only modest divergence appearing at higher job indices. This closeness is by design, as the firm-level elasticity ε_δ is recalibrated to match this target. The lesson, however, is not that heterogeneity is unimportant — it is that the worker churn ladder, measured as a proportional slope across transitions, is insufficiently informative on its own to distinguish the two mechanisms. The following figures show that the model without

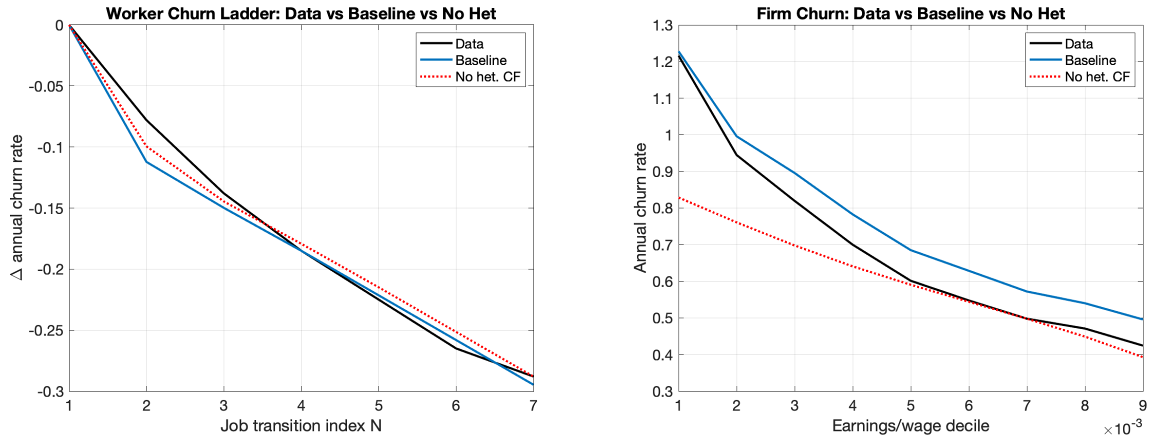

Figure 10: LEFT: The change in annual churn rate of the employer relative to the initial job across job ladder transitions in the data, the calibrated baseline model, and the no-heterogeneity counterfactual. The average churn rate of the initial job is normalized to 0. RIGHT: The mean annual churn rate at each decile of the firm earnings distribution in the data, the calibrated baseline model, and the no-heterogeneity counterfactual.

heterogeneity achieves this partial fit only by failing along every other margin.

The firm-level churn gradient reveals the first failure without heterogeneity. The right panel of Figure 10 plots the annual churn rate against firm wage decile for the data, baseline, and no-het counterfactual. While the baseline closely tracks the data across the full distribution, the no-het counterfactual understates churn substantially at the bottom of the wage distribution, converging toward the data only at the top. The shortfall at the bottom is mechanical: in the baseline, low-wage firms accumulate high-δ workers who separate frequently to non-employment, amplifying gross worker flows beyond what poaching alone can generate. Without that compositional channel, the firm-level elasticity cannot push churn at low-wage firms up to observed levels, even when recalibrated to target the worker ladder slope.

The distribution of job tenure makes the failure most direct, and connects most clearly to the tests developed in Section 3.2. Figure 11 plots the fraction of workers with tenure under one year across firm wage deciles. The baseline model tracks the data closely, while the no-het counterfactual overshoots uniformly — predicting too many short-tenure matches at every point in the wage distribution, with the excess concentrated at lower wages. The reason is

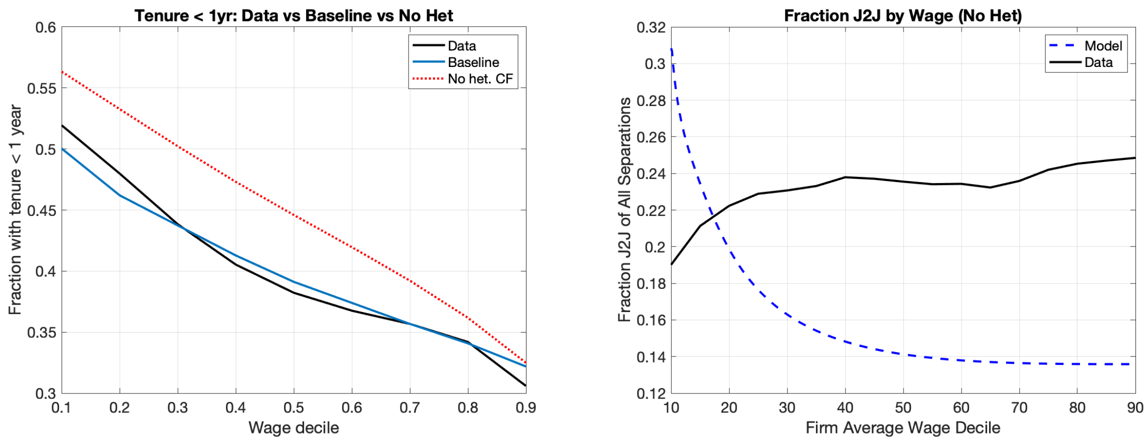

Figure 11: LEFT: The fraction of workers with less than one year of tenure at each decile of the firm wage distribution in the data, the calibrated baseline model, and the no-heterogeneity counterfactual. The no-heterogeneity counterfactual overpredicts short-tenure matches throughout the distribution, with the excess most pronounced at lower wages. RIGHT: The fraction of all separations that are job-to-job transitions at each decile of the firm earnings distribution in the data and the no-heterogeneity counterfactual.

precisely the tension identified in Section 3.2: with a single separation rate within each wage bin, churn and tenure are linked one-for-one, leaving no degrees of freedom to match both simultaneously. The baseline breaks this link through the mixture of high- and low-δ types at each wage level, whose differing survival rates generate the excess dispersion in tenure the data require. Without this mixture, the firm-level policy must set higher separation rates at the bottom to approach the observed churn gradient, which in turn inflates the predicted short-tenure fraction beyond what we observe.

The composition of separations provides the most decisive evidence against the no-het model. Figure 11 (RIGHT) plots the fraction of all separations that are job-to-job, by firm wage decile, comparing the no-het model to the data. In the data, the J2J share of separations is roughly flat across the wage distribution, even rising modestly at the top. The no-het model predicts the opposite: a steep monotonic decline from well above the data at the bottom to well below it at the top. This reversal reflects the fundamental reallocation that worker heterogeneity performs in the baseline. Elevated churn at low-wage firms is driven primarily by non-employment separations among the concentrated high-δ workers, not by poaching. Without that compositional force, the no-het model must attribute excess churn at the bottom to J2J transitions — overstating that margin dramatically and distorting the ratio of EN to EE separations in a direction that is directly at odds with the data.

Taken together, these figures establish that worker heterogeneity is not a peripheral calibration device. Removing it produces a model that can approximately replicate the slope of the worker churn ladder but fails on every other margin: the level of firm churn at the bottom of the distribution, the tenure distribution within wage bins, and the composition of separations across the ladder. The baseline model's success across all four dimensions reflects the structural role of endogenous sorting: as low-δ workers accumulate at high-wage firms through longer employment spells, they simultaneously compress firm churn at the top, generate excess dispersion in tenure within wage bins, and keep the J2J share of separations roughly constant across the wage distribution — exactly the joint pattern the data exhibit.

5.4 Measuring labor supply elasticity and monopsony

An important application of our churn ladder is how it informs measures of labor market power. In most job ladder models, poaching is the driving feature that makes the labor supply faced by firms dependent on wages. Labor supply depends on wages both in how firms use them to retain workers and in how they use them to recruit from other firms. Further, labor supply elasticity directly informs the firms' monopsonistic markdown. For a higher elasticity of supply, firms enjoy less market power and must pay more competitive wages. This is parsimoniously demonstrated in Manning (2003) and Bassier et al. (2021), in which the firm's separation elasticity with respect to the wage is the key element to measure. The crucial insight to this literature from our work is also the finding that allows us to match the churn ladder: separations to non-employment vary systematically with wages so poaching is not the only factor moving the elasticity with wages.

To derive the separation elasticity with the notation in Manning (2003), the rate of outflow from firms is a function of the wage level and identical to our measure of the churn rate in Equation 4.2, $\delta + \lambda(1 - F(w))$. Similarly, we can derive the recruiting inflow to be

$\lambda u + \lambda G(w)(1-u)$. In steady state, at every w, $L(w) = \frac{\lambda u + \lambda G(w)(1-u)}{\delta + \lambda(1-F(w))}$. In this simpler model, akin to the one presented in Manning (2003), we can derive a labor supply elasticity which directly relates to the monopsonistic markdown. With δ constant in w, it is $\frac{dL}{dw}\frac{w}{L} = \frac{2\lambda f(w) w}{\delta + \lambda(1-F(w))}$, but once $\delta'(w) \neq 0$, it becomes $\frac{w\delta'(w)}{\delta(w)}\frac{\lambda(1-F(w))-\delta(w)}{\lambda(1-F(w))+\delta(w)} + \frac{2\lambda f(w) w}{\delta(w)+\lambda(1-F(w))}$.[12] The last term is the same in both cases, but the first term when $\delta'(w) \neq 0$ is a positive number. It is the elasticity of δ with respect to wages, a negative number, times a fraction that is -1 for the maximum wage and a smaller negative number below that as long as λ is not larger than $\delta(\underline{w})$. Because $\underline{w}$ is where $\delta(w)$ is at its largest, this is unlikely.

With our full model from Section 4.1, we can also show that $\delta'(w) < 0$ makes the overall labor supply elasticity larger at a given wage and thereby reduces monopsonistic markdowns. Here the $\delta(w)$ might come from composition or from firm effects. Our labor supply with respect to w is

$$L(w) = \frac{\lambda_u u + (\lambda_r + \lambda_e G(w))(1-u)}{\delta(w) + \lambda_r + \lambda_e(1-F(w))} \tag{5.1}$$

Again, solving first for the elasticity of labor were $\delta'(w) = 0$, we would get[13]

$$w\left(\frac{\lambda_e(1-u)g(w)}{\lambda_u u + (\lambda_r + \lambda_e G(w))(1-u)} + \frac{\lambda_e f(w)}{\delta + \lambda_r + \lambda_e(1-F(w)))}\right). \tag{5.2}$$

If, however, we calculate it in Equation 5.1 with $\delta'(w) < 0$, then we get for $\frac{dL}{dw}\frac{w}{L}$

$$\frac{dL}{dw}\frac{w}{L} = w\left(\frac{\lambda_e(1-u)g(w)}{\lambda_u u + (\lambda_r + \lambda_e G(w))(1-u)} + \frac{\lambda_e f(w) - \delta'(w)}{(\delta(w) + \lambda_r + \lambda_e(1-F(w)))}\right). \tag{5.3}$$

Notice the only difference between the two is that the second term of Equation 5.3 has a $-\delta'(w)$ in its numerator. Because $\delta'(w) < 0$, this is a positive term, the labor supply elasticity faced by the firm is now larger with respect to wages.

To show this graphically, we compute the labor supply elasticity in our calibrated model from Equation 5.2 and from Equation 5.3.

[12] Our model in Section 4.1 nests the latter case with $\delta'(w) \neq 0$ if $\lambda_u = \lambda_e$ and $\lambda_r = 0$.
[13] All of our derivations are in Appendix Section D.

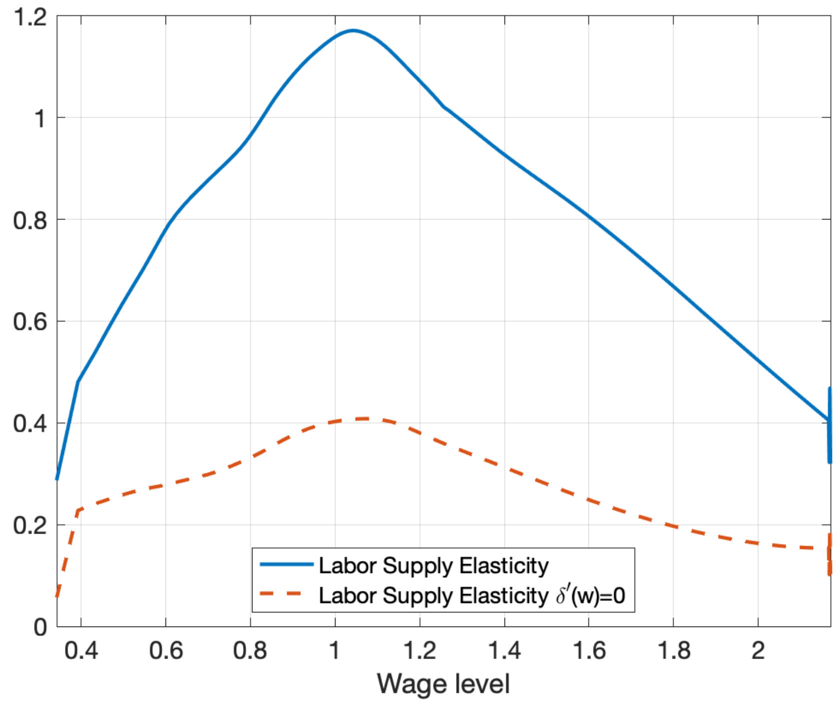

Figure 12: Labor Supply Elasticity

Figure 12 plots the labor supply elasticity faced by the firm across the wage distribution under two assumptions: our baseline calibrated model, in which the separation rate declines with wages ($\delta'(w) < 0$), and a restricted version that holds δ constant at its mean value. The gap between the two curves is striking. Across virtually the entire wage distribution, the baseline elasticity is roughly two to three times larger than the constant-δ case, with the difference most pronounced in the middle of the distribution where both the density of offers and the slope of $\delta(w)$ are large. This gap directly translates into lower implied monopsony markdowns: firms face a more wage-elastic labor supply than a model based solely on poaching flows would suggest. The shape of the baseline elasticity is also informative. It rises steeply from the bottom of the wage distribution, peaks near the median wage, and then declines toward the top. The initial rise reflects the increasing density of the offer distribution and the steepest portion of the $\delta(w)$ gradient, where both the poaching margin and the nonemployment separation margin respond most sharply to wages. The decline at the top of the distribution reflects the diminishing mass of potential poachers–$1 - F(w)$ approaches zero–and the flattening of $\delta(w)$ as the composition of workers stabilizes at nearly all low-separation types. The constant-δ curve, by contrast, rises monotonically and modestly, driven entirely by the offer distribution, and misses both the peak and the subsequent decline.

The practical implication is that studies relying exclusively on job-to-job flows to estimate the separation elasticity–and thereby infer firms' wage markdowns–will systematically understate how responsive total worker retention is to wages. In our calibrated model, the omitted channel is large: nonemployment separations account for the majority of the gap between the two elasticity curves throughout the distribution. This suggests that empirical estimates of monopsony power derived from event-study or quasi-experimental designs that focus on poaching flows should be interpreted as lower bounds on the true labor supply elasticity, and that incorporating wage-sensitive separations to nonemployment would meaningfully reduce inferred markdowns, bringing them closer to the more competitive benchmarks found in reduced-form studies.

5.5 Discussion: different dimensions for measuring the job ladder

Theory suggests that on-the-job search creates a job ladder and thereby shapes the relationship between firm size, wages, and churn. Here, we find strong support for this mechanism: wages and churn vary across workers and firms in a manner broadly consistent with the canonical job ladder model. The main discrepancy is quantitative, as the magnitudes are off. The theoretical relationship is driven by high-wage firms poaching from low-wage firms, increasing their churn. But this poaching mechanism yields a familiar finding in the search literature: the dispersion we observe in the data is much greater than that predicted by frictional labor market models. Nevertheless, frictions are still essential for generating the dispersion we observe because they facilitate sorting sorting across firms by wage and across workers by separation risk. This sorting mechanism, together with on-the-job search, helps to explain the empirical patterns we document.

The importance of nonemployment separations informs how job ladder models rank firms. On the one hand, job ladders create a natural ranking of firms and we suggest another dimension by which they can be measured. On the other hand, our results suggest caution in selecting the appropriate dimension to measure.

In foundational work like Bagger and Lentz (2018) or Sorkin (2018), firm rankings can be taken from either worker transition patterns or firm hiring patterns. Our proposed ordering of firms' separation patterns is also consistent in rank, but not magnitude with these approaches. Firms at the bottom of the ladder are both hiring less from other firms and also losing more workers, but the variation in these losses over the wage distribution is much larger than would be implied by poaching alone. That is, while a firm's poaching probability aligns with its churn rank, the overall churn ladder is cardinally much steeper than what poaching flows alone imply.

This extra steepness is important for papers that try to match the differences in the pattern of firm-level flows in the cross-section. Job-to-job transitions alone cannot create the heterogeneity in firm-level employee flows nor the stark contrast in churn rates between low- and high-wage firms. The literature working on the foundations of Burdett and Mortensen (1998) or Postel-Vinay and Robin (2002) implies cross-firm differences in churn that are too low and insufficiently correlated with wages unless explicitly incorporating nonemployment separation heterogeneity as we did with our two mechanisms.

This exercise also has lessons for other methods of measuring rank, looking at firm characteristics like wage, productivity or size (see, e.g. Moscarini and Postel-Vinay (2012) or Haltiwanger et al. (2018)). Again, the position on the churn ladder does not change the rank order, but it provides a more direct measure than size. In a model, higher rank firms poach from lower rank firms and thereby become larger. However, size can also be affected by other factors, e.g. production functions and the Coaseian contracting environment. So rather than looking at this second order implicate, we suggest looking at churn itself.

Our model included both firm and worker heterogeneity in unemployment separations and distinguished the contribution of each, which sheds light on several important lines of research that feature firm or worker separation heterogeneity. Of course, ours was a fairly parsimonious model so the micro foundations of firm-level differences are not settled in this paper.

Firm-level differences in separation rates have been crucial to the "slippery" job ladder mechanisms that try to explain persistent earnings losses after job loss. This literature uses a steep $e^{-\varepsilon_\delta w}$ to explain the experience of displaced workers, watching them experience repeated unemployment spells (see Pinheiro and Visschers (2015) or Jarosch (2023)). The churn ladder offers a complementary perspective on this phenomenon: high separation rate firms at the bottom rungs of the job ladder, those that make earnings losses more persistent for separated workers, are one ingredient to the churn ladder. But the degree to which a steep gradient of $\delta(w)$ is because of firms themselves, $e^{-\varepsilon_\delta w}$, or worker sorting, $\mu(w)$, informs the literature's notions of job ladder slipperiness. Our estimates suggest a sizeable portion of the high separation rates at these firms is because of the concentration of high separation workers they have. Thus slow earnings recovery is not only because of the firms displaced workers match with, but also because these workers were unlikely to climb the ladder in the first place.

Unobservable worker-level heterogeneity is a key driver for our churn ladder and shapes the cross-sectional distribution of firms. Similar worker-side differences in job separation rates have also been central in the recent discussion of heterogeneous unemployment rates. This work has, with worker-side data, shown large differences in the separation and unemployment rates across workers (see Hall and Kudlyak (2019), Ahn et al. (2023), or Gregory et al. (2021)) to better understand business cycle fluctuations in unemployment.Again, the churn ladder provides another perspective on the role of worker heterogeneity. Not only does the high separation worker drive the unemployment rate, it also elevates worker churn rates at the bottom of the job ladder. In this context, because of endogenous sorting our model would predict that high separation firms play an analogously crucial role to recessionary increases in unemployment because of their concentration of these cyclically sensitive workers. Leaving this postulate for future work, our churn ladder also introduces evidence that corroborates Borovičková and Macaluso (2023), in which worker-heterogeneity is crucial to individuals' earnings dynamics.

6 Conclusions

In this paper, we combine new evidence on the empirical shape of the churn ladder from both worker- and firm-side observations. This evidence is qualitatively consistent with a broad class of models, that the churn ladder exists in tandem with the wage ladder. We estimate the firm-side negative relationship between earnings and churn and document their significant dispersion in the cross-section. Estimating the slope of both wage and churn over the job ladder gives a quantitative relationship that can be directly compared to calibrated workhorse models of on-the-job search.

The failure of the job ladder model to generate the dispersion in turnover that we see in the data is informative of an important missing mechanism. To match the turnover rates at the top of the firm ranking, firms must experience far fewer separations to both poaching and nonemployment. We demonstrate that there is substantial heterogeneity in separation rates across employers within our data, and more importantly, separations to nonemployment are highly correlated with the firm's wage rank. Heterogeneity in worker separations and their implied sorting is consistent with a growing literature identifying latent heterogeneity in workers' attachment to employment and the reconciliation of worker flows. Our finding elaborates on this literature: heterogeneity in workers' separation rates seems to play a key role in the distribution of gross flows across the firm distribution as well.

Our paper confirms the job ladder model's relationship between turnover and earnings while also highlighting the need for more heterogeneity. What we observe empirically is excessive firm-level heterogeneity in the amount of turnover, which, while it qualitatively goes in the same direction as the workhorse job ladder model would predict, is much more dispersed. Our theory, however, shows that this firm-level heterogeneity can be created by worker-side heterogeneity. We think this is an important insight: that these firm-level patterns are equally shaped by worker-level differences and sorting.

References

Ahn, H. J., B. Hobijn, and A. Şahin (2023): "The Dual U.S. Labor Market Uncovered," NBER Working Papers 31241, National Bureau of Economic Research, Inc.

Bagger, J. and R. Lentz (2018): "An Empirical Model of Wage Dispersion with Sorting," *The Review of Economic Studies*, 86, 153–190.

Baksy, A., D. Caratelli, and N. Engbom (2024): "The Long-term Decline of the US Job Ladder," Working paper.

Barlevy, G. (2008): "Identification of Search Models using Record Statistics," *The Review of Economic Studies*, 75, 29–64.

Bassier, I., A. Dube, and S. Naidu (2021): "Monopsony in Movers: The Elasticity of Labor Supply to Firm Wage Policies," *Journal of Human Resources*.

Borovičková, K. and C. Macaluso (2023): "Heterogeneous Job Ladders," Working paper.

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

Elsby, M., R. Michaels, and D. Ratner (2017): "Vacancy Chains," 2017 Meeting Papers 888, Society for Economic Dynamics.

Gottfries, A. and C. Teulings (2023): "Returns to on-the-job search and wage dispersion," *Labour Economics*, 80, 102292.

Gregory, V., G. Menzio, and D. G. Wiczer (2021): "The Alpha Beta Gamma of the Labor Market," NBER Working Papers 28663, National Bureau of Economic Research, Inc.

Hahn, J. K., H. R. Hyatt, H. P. Janicki, and S. R. Tibbets (2017): "Job-to-Job Flows and Earnings Growth," *American Economic Review*, 107, 358–63.

Hall, R. E. and M. Kudlyak (2019): "Job-Finding and Job-Losing: A Comprehensive Model of Heterogeneous Individual Labor-Market Dynamics," NBER Working Papers 25625, National Bureau of Economic Research, Inc.

Haltiwanger, J. C., H. R. Hyatt, L. B. Kahn, and E. McEntarfer (2018): "Cyclical job ladders by firm size and firm wage," *American Economic Journal: Macroeconomics*, 10, 52–85.

Hornstein, A., P. Krusell, and G. L. Violante (2011): "Frictional Wage Dispersion in Search Models: A Quantitative Assessment," *American Economic Review*, 101, 2873–98.

Hyatt, H. R. and J. R. Spletzer (2017): "The recent decline of single quarter jobs," *Labour Economics*, 46, 166–176.

Jarosch, G. (2023): "Searching for Job Security and the Consequences of Job Loss," *Econometrica*, 91, 903–942.

Jolivet, G., F. Postel-Vinay, and J.-M. Robin (2006): "The empirical content of the job search model: Labor mobility and wage distributions in Europe and the US," Université Paris1 Panthéon-Sorbonne (Post-Print and Working Papers) hal-00279066, HAL.

Karahan, F., S. Ozkan, and J. Song (2019): "Anatomy of Lifetime Earnings Inequality: Heterogeneity in Job Ladder Risk vs. Human Capital," Staff Reports 908, Federal Reserve Bank of New York.

Lamadon, T., J. Lise, C. Meghir, and J.-M. Robin (2024): "Labor Market Matching, Wages, and Amenities," Tech. rep., National Bureau of Economic Research.

Lentz, R., S. Piyapromdee, and J.-M. Robin (2023): "The Anatomy of Sorting—Evidence From Danish Data," *Econometrica*, 91, 2409–2455.

Manning, A. (2003): *Monopsony in Motion: Imperfect Competition in Labor Markets*, Princeton University Press.

Moscarini, G. and F. Postel-Vinay (2008): "The Timing of Labor Market Expansions: New Facts and a New Hypothesis," *NBER Macroeconomics Annual*, 23, 1–52.

——— (2012): "The contribution of large and small employers to job creation in times of high and low unemployment," *American Economic Review*, 102, 2509–2539.

——— (2017): "The Relative Power of Employment-to-Employment Reallocation and Unemployment Exits in Predicting Wage Growth," *American Economic Review*, 107, 364–68.

Pinheiro, R. and L. Visschers (2015): "Unemployment risk and wage differentials," *Journal of Economic Theory*, 157, 397–424.

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

Sorkin, I. (2018): "Ranking Firms Using Revealed Preference," *The Quarterly Journal of Economics*, 133, 1331–1393.

Topel, R. H. and M. P. Ward (1992): "Job Mobility and the Careers of Young Men," *The Quarterly Journal of Economics*, 107, 439–479.

Wolpin, K. I. (1992): "The determinants of black-white differences in early employment careers: Search, layoffs, quits, and endogenous wage growth," *Journal of Political Economy*, 100, 535–560.

A Appendix: additional descriptive statistics on the churn ladder

A.1 Summary statistics by size and industry

In this section we present a few more facts about the distribution of these joint churn and wage ladders. There is a great amount of heterogeneity across firms in both churn and average wages, and most of it exists within observable characteristics.

First, we break down the wage and churn distributions across firm size and firm size changes. We show in Table 2 that annual churn is decreasing in size and lowest among firms that do not change their size.

Table 2: Mean and variance of churn by employment size and growth

	Mean Churn	Var. of churn within group
Small (< 50)	0.821	2.465
Med (50-1000)	0.794	1.358
Large (1000+)	0.616	0.706
Shrinking ($< -2\%$)	0.814	1.975
Stable $[-2\%, 2\%]$	0.567	0.508
Growing ($> 2\%$)	0.875	2.78

We also note that there is considerable variation in churn across industries. Table 3 lists the mean and variance of the annual churn rate for firms within each sector. Note that there is considerable variation across industries in both the mean and variance of churn rates. However, this variation in churn across industries appears to be largely driven by differences in wages. In Figure 13, we plot sector-level means of churn rates against sector-level means of firms' average quarterly earnings. We get a familiar negative relationship between wages and churn and similar dispersion in churn rates across the wage distribution, just as we see in the unconditional relationship in Figure 1.

Churn vs. Wage by Sector

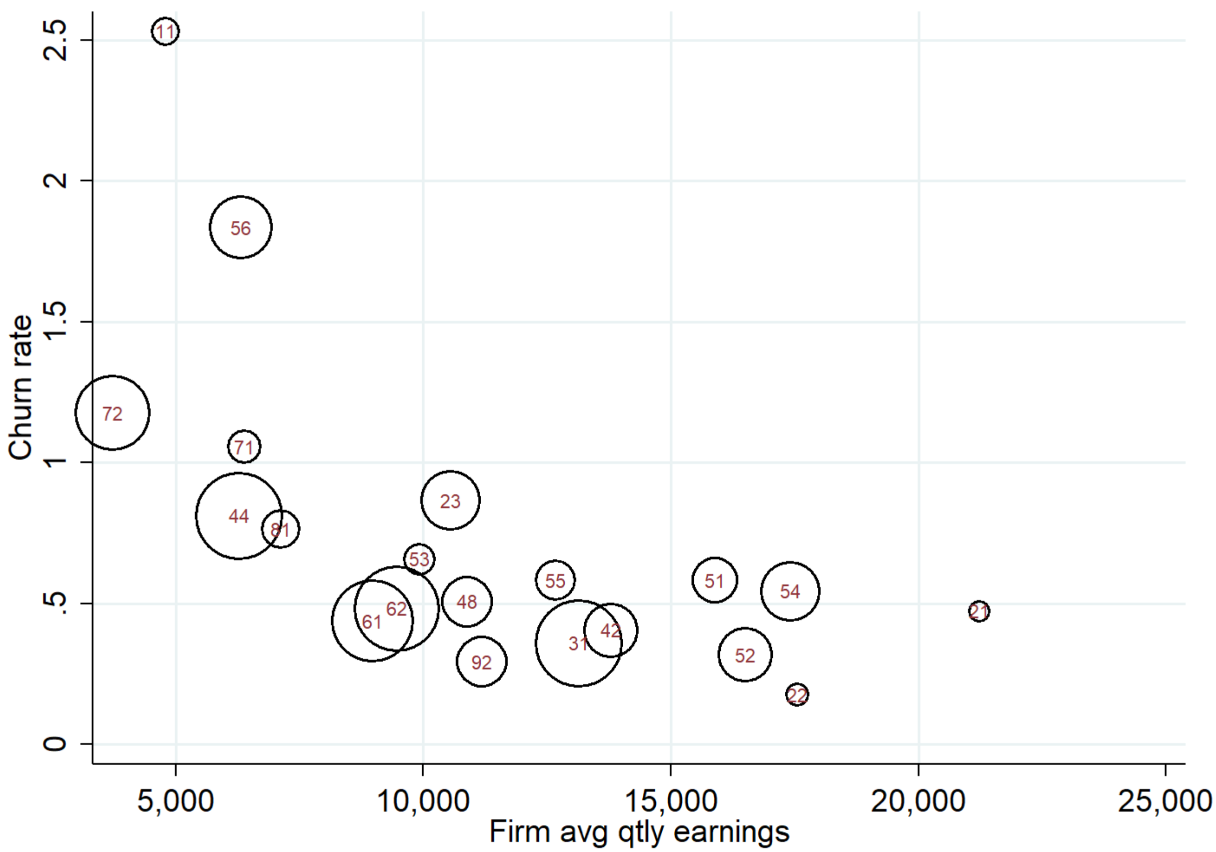

Figure 13: Employment-weighted mean of annual churn rate of firms within each sector, plotted against the employment-weighted mean of firms' average quarterly earnings. Bubble size represents the employment weight of the sector. Sector labeled 31 represents Manufacturing (31-33), Sector 44 represents Retail Trade (44-45), and Sector 48 represents Transportation and Warehousing (48-49).

Table 3: Mean and variance of churn by industry sector.

	Mean Churn	Var. of churn within group
NAICS 11: Ag., Forestry, Fishing, Hunting	2.443	21.76
NAICS 21: Mining, Quarrying, Oil & Gas Extr.	0.685	0.716
NAICS 22: Utilities	0.268	0.350
NAICS 23: Construction	0.980	1.239
NAICS 31-33: Manufacturing	0.515	0.647
NAICS 42: Wholesale Trade	0.458	0.576
NAICS 44-45: Retail Trade	0.812	1.202
NAICS 48-49: Transport. & Warehousing	0.728	1.099
NAICS 51: Information	0.597	8.649
NAICS 52: Finance & Insurance	0.400	1.148
NAICS 53: Real Estate & Rental & Leasing	0.662	1.063
NAICS 54: Prof., Sci., & Tech. Services	0.520	1.310
NAICS 55: Mgmt. of Companies & Enterprises	0.581	5.612
NAICS 56: Admin. & Supp. etc. Services	1.347	5.986
NAICS 61: Education	0.663	2.515
NAICS 62: Health Care & Social Assistance	0.620	1.218
NAICS 71: Arts, Entertainment, & Recreation	1.311	5.383
NAICS 72: Accommodation & Food Services	1.266	1.250
NAICS 81: Other Services	0.765	1.192
NAICS 99: Public Administration	0.484	0.653

B Appendix: Robustness of the worker churn ladder

In this section, we plot the worker-level churn ladder across different groups. The churn ladder is remarkably robust across observable characteristics of the worker. The decline in churn across job indexes within workers' continuous employment spells is consistent in slope across sex, race, education level, and age. There are significant variations in the levels of churn over these employment cycles, but the decline across job transitions is remarkably consistent.

Of note as well is that the slope of the worker churn ladder is very similar across workers' lifetime income profiles. In Figure 18, we plot the churn ladder for each tercile of lifetime earnings, conditioning on birth cohort. High- and low lifetime earnings workers experience similar churn ladders.

We also note that the worker-level churn ladder holds across firm characteristics such as size and growth, despite the variation in churn levels in these firms that we found in Tables 2 and 3. While we see significant differences in the initial levels of churn across sectors, the worker-level coefficients at each job index within sectors displays very similar negative slopes in terms of percentage changes.

Churn Ladder: Sex

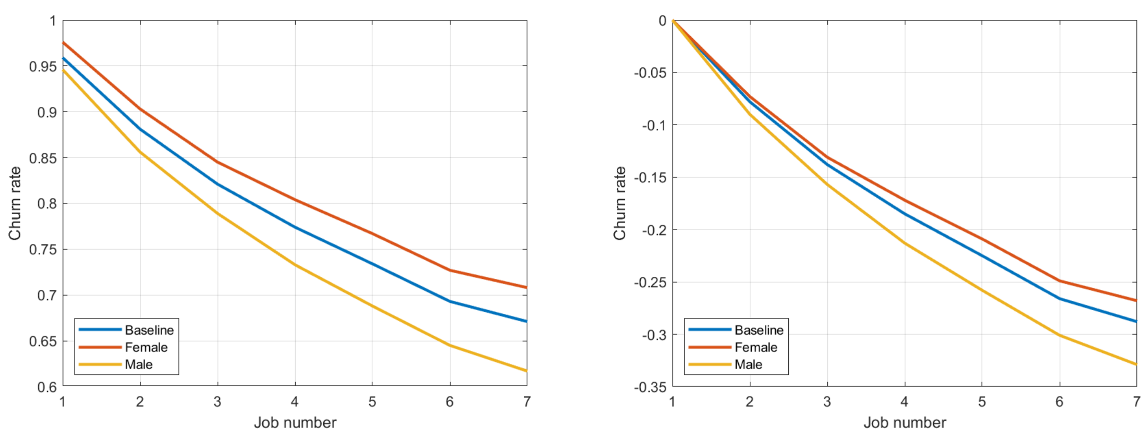

Figure 14: Worker-level churn ladder estimates in levels (LEFT) and relative to initial job (RIGHT).

Churn Ladder: Race

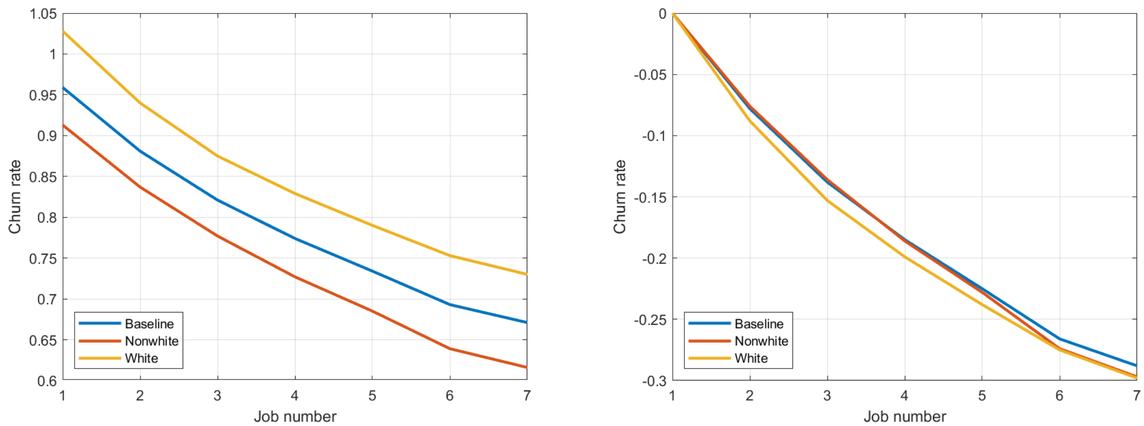

Figure 15: Worker-level churn ladder estimates in levels (LEFT) and relative to initial job (RIGHT).

Churn Ladder: Age

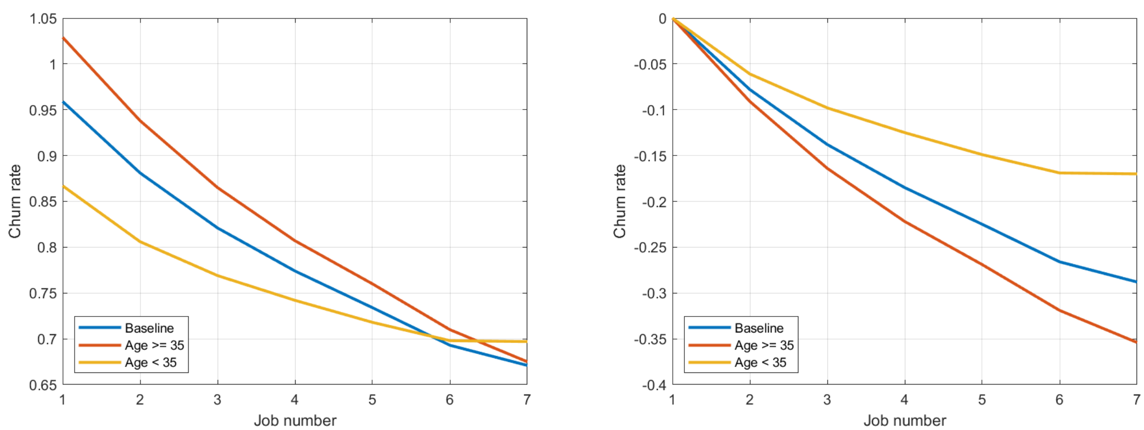

Figure 16: Worker-level churn ladder estimates in levels (LEFT) and relative to initial job (RIGHT).

Churn Ladder: Education

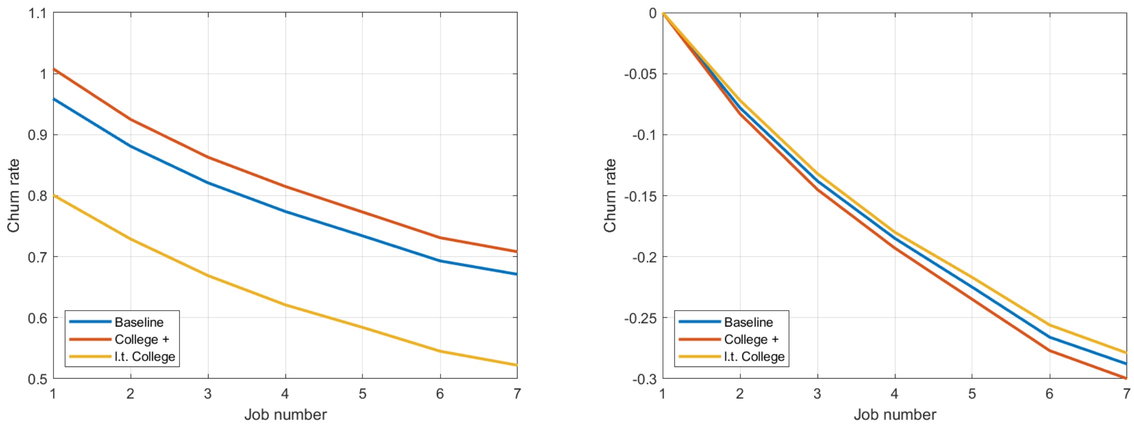

Figure 17: Worker-level churn ladder estimates in levels (LEFT) and relative to initial job (RIGHT).

Churn Ladder: Lifetime Income Terciles (cond. on cohort)

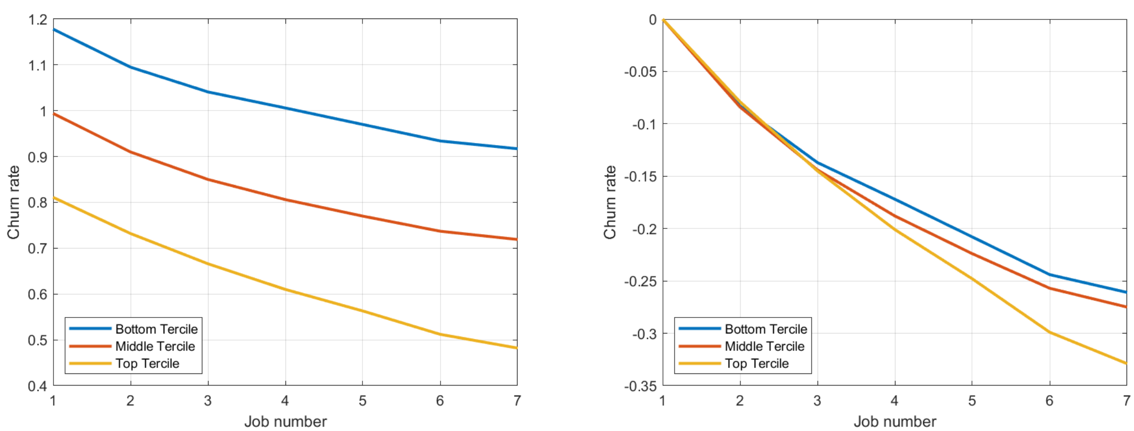

Figure 18: Worker-level churn ladder estimates in levels (LEFT) and relative to initial job (RIGHT).

Churn Ladder: Firm Size

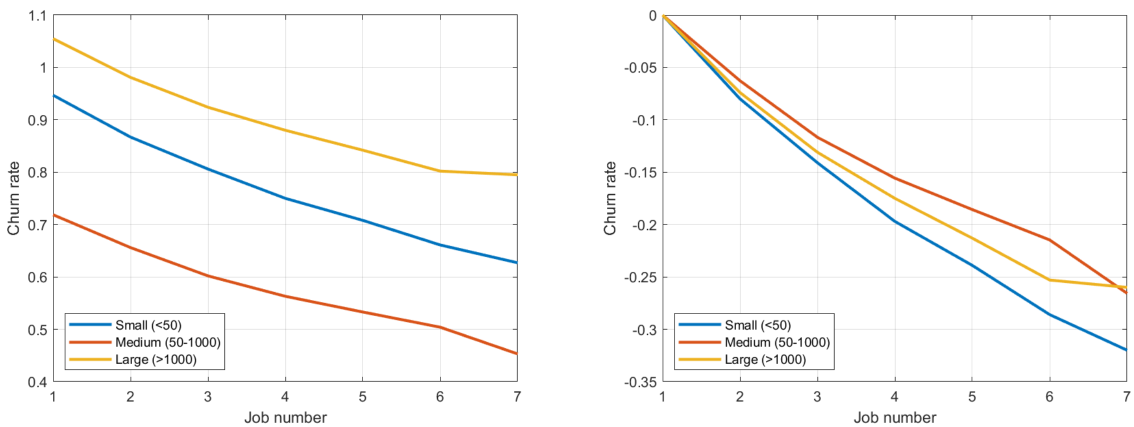

Figure 19: Worker-level churn ladder estimates in levels (LEFT) and relative to initial job (RIGHT).

Churn Ladder: Firm Growth

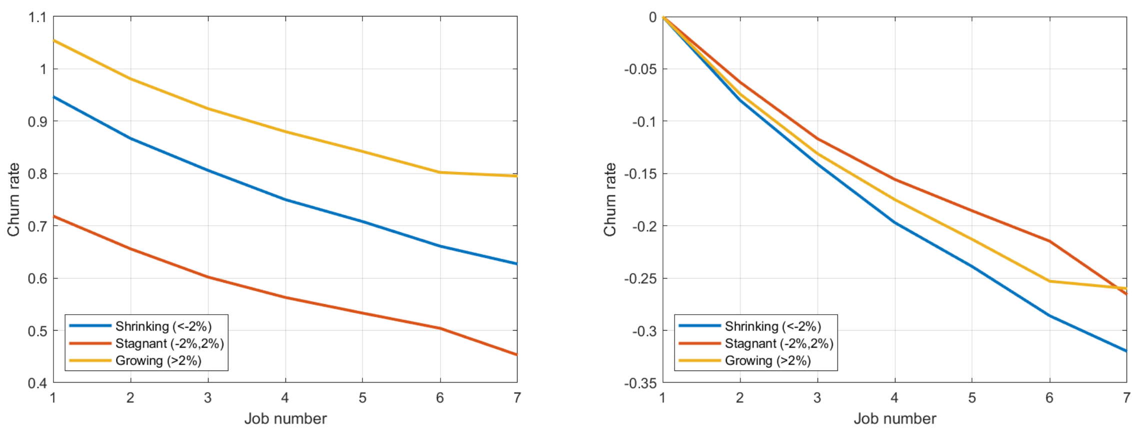

Figure 20: Worker-level churn ladder estimates in levels (LEFT) and relative to initial job (RIGHT).

Churn Ladder: By Sectors

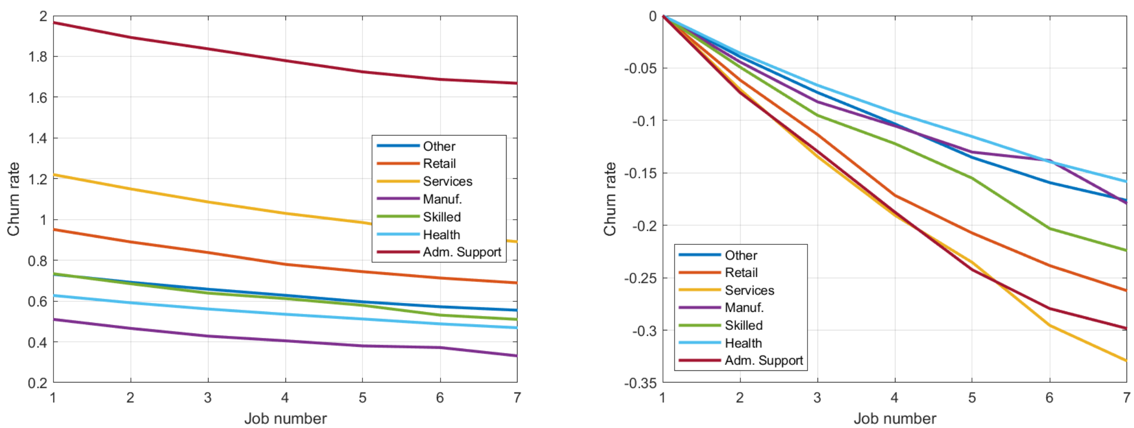

Figure 21: Worker-level churn ladder estimates in levels (LEFT) and relative to initial job (RIGHT).

C Appendix: deriving the identification with discrete types

C.1 Mapping model to data

Suppose wages are defined on a discrete interval, now with discrete densities $\tilde{h}$. Then for discrete wage value w_k from range $w \in \{w_1, \dots, w_W\}$ and $\delta_j \in \{\delta_1, \dots, \delta_\Delta\}$ we have the following equation (4.6):

$$\sum_{i=1}^{k}\sum_{j=1}^{\Delta}(\delta_j e^{\varepsilon\delta w_i} + (\lambda_r + \lambda_e)(1 - F(w_k)))(1 - u_j)h(w_i, \delta_j) =$$

$$\lambda_u uF(w_k) + F(w_k)\sum_{i=k+1}^{W}\sum_{j=1}^{\Delta}\lambda_r(1 - u_j)h(w_i, \delta_j)$$

Where $\delta(w_k) = \sum_{j=1}^{\Delta}\delta_j\tilde{h}(w_k, \delta_j)$, that is the effective separation rate for a firm at wage W_k is the worker-weighted average of separation rates δ_j for each j type of worker at wage w_k. Suppose there are two types, δ_1, δ_2 with underlying probability $p_1, 1 - p_1$. we can write (4.6) as:

With two types, δ_1, δ_2 with underlying probability $p_1, 1-p_1$ our identification comes from solving a fairly straightforward system.

We can write (4.6) as:

$$\begin{aligned}
&\sum_{i=1}^{k}(\delta_1 e^{\varepsilon\delta w_k} + \lambda_r(1 - F(w_k)) + \lambda_e(1 - F(w_k)))(1 - u_1)h(w_i, \delta_1) + \\
&\quad(\delta_2 e^{\varepsilon\delta w_k} + \lambda_r(1 - F(w_k)) + \lambda_e(1 - F(w_k)))(1 - u_2)h(w_i, \delta_2) = \\
&\quad\lambda_u uF(w_k) + F(w_k)\sum_{i=k+1}^{W}\lambda_r(1 - u_1)h(w_i, \delta_1) + F(w_k)\sum_{i=k+1}^{W}\lambda_r(1 - u_2)h(w_i, \delta_2)\ .
\end{aligned}$$

Given a $G(w)$ and $\delta(w)$, our joint densities $h(w_i, \delta_j)$ are defined by the $2 \times N$ conditions:

$$\begin{aligned}
\delta(w_i) &= \left(\delta_1\frac{h(w_i, \delta_1)}{g(w_i)} + \delta_2\frac{h(w_i, \delta_2)}{g(w_i)}\right)e^{\varepsilon\delta w_i} \\
g(w_i) &= h(w_i, \delta_1) + h(w_i, \delta_2) \\
&\text{for } i \in \{1, ..., N\}
\end{aligned}$$

If we fixed the types δ_j, then to solve for the masses p_1 and $(1 - p_1)$ of each worker type $j \in \{1, 2\}$, we can use the relationship between the aggregate unemployment rate u and unemployment rate for each type, u_j and their steady-state relationship to inflows and outflows.

Now to get the masses p_1 and $(1-p_1)$ of each worker type $j \in \{1,2\}$, we can use the relationship between the aggregate unemployment rate u and unemployment rate for each type, u_j and their steady-state relationship to inflows and outflows.

$$u = \sum_j u_j p_j = u_1 p_1 + u_2(1-p_1) = \frac{\delta_1}{\delta_1 + \lambda_u} p_1 + \frac{\delta_2}{\delta_2 + \lambda_u}(1-p_1)$$

$$u = \left(\frac{\delta_1}{\delta_1 + \lambda_u} - \frac{\delta_2}{\delta_2 + \lambda_u} \right) p_1 + \frac{\delta_2}{\delta_2 + \lambda_u}$$

And it should be that for the employed, the following holds:

$$\frac{p_j(1-u_j)}{1-u} = \sum_{i=1}^{N} h(w_i, \delta_j) \text{ for } j \in \{1,2\}$$

D Labor supply elasticity derivations

The firms size is determined by flow balances:

$$\delta(w) + \lambda_r + \lambda_e(1 - F(w))L(w) = \lambda_u u + (\lambda_r + \lambda_e G(w))(1 - u)$$

Solving for $L(w)$ gives

$$L(w) = \frac{\lambda_u u + (\lambda_r + \lambda_e G(w))(1 - u)}{\delta(w) + \lambda_r + \lambda_e(1 - F(w))}$$

Supposing that $\delta'(w) = 0$, let's find the elasticity $\varepsilon_L(w)$.

$$\frac{\partial L(w)}{\partial w}\frac{w}{L} = \frac{(1-u)g(w)\lambda_e(\delta + \lambda_r + \lambda_e(1 - F(w))) + (\lambda_u u + (\lambda_r + \lambda_e G(w))(1-u))\lambda_e f(w)}{(\delta + \lambda_r + \lambda_e(1 - F(w)))^2}\frac{L}{w} \tag{D.1}$$

$$= w\frac{(1-u)g(w)\lambda_e(\delta + \lambda_r + \lambda_e(1 - F(w))) + (\lambda_u u + (\lambda_r + \lambda_e G(w))(1-u))\lambda_e f(w)}{(\delta(w) + \lambda_r + \lambda_e(1 - F(w)))^2}\frac{\delta + \lambda_r + \lambda_e(1 - F(w))}{\lambda_u u + (\lambda_r + \lambda_e G(w))(1-u)} \tag{D.2}$$

$$= w\frac{(1-u)g(w)\lambda_e(\delta + \lambda_r + \lambda_e(1 - F(w)))}{(\delta + \lambda_r + \lambda_e(1 - F(w)))(\lambda_u u + (\lambda_r + \lambda_e G(w))(1-u))} + w\frac{(\lambda_u u + (\lambda_r + \lambda_e G(w))(1-u))\lambda_e f(w)}{(\delta + \lambda_r + \lambda_e(1 - F(w)))(\lambda_u u + (\lambda_r + \lambda_e G(w))(1-u))} \tag{D.3}$$

$$= w\frac{(1-u)g(w)\lambda_e(\delta + \lambda_r + \lambda_e(1 - F(w)))}{(\delta + \lambda_r + \lambda_e(1 - F(w)))(\lambda_u u + (\lambda_r + \lambda_e G(w))(1-u))} + w\frac{\lambda_e f(w)}{\delta + \lambda_r + \lambda_e(1 - F(w)))} \tag{D.4}$$

$$= w\frac{(1-u)g(w)\lambda_e A}{AB} + w\frac{\lambda_e f(w)}{\delta + \lambda_r + \lambda_e(1 - F(w)))} \tag{D.5}$$

$$= w\frac{(1-u)g(w)\lambda_e}{\lambda_u u + (\lambda_r + \lambda_e G(w))(1-u)} + w\frac{\lambda_e f(w)}{\delta + \lambda_r + \lambda_e(1 - F(w)))} \tag{D.6}$$

$$= w\lambda_e\left(\frac{(1-u)g(w)}{\lambda_u u + (\lambda_r + \lambda_e G(w))(1-u)} + \frac{f(w)}{\delta + \lambda_r + \lambda_e(1 - F(w)))}\right) \tag{D.7}$$

$$\tag{D.8}$$

Now allowing $\delta'(w) \neq 0$

$$\frac{\partial L(w)}{\partial w}\frac{w}{L} = \frac{(1-u)g(w)\lambda_e(\delta(w)+\lambda_r+\lambda_e(1-F(w))) - (\lambda_u u + (\lambda_r+\lambda_e G(w))(1-u))(\delta'(w)-\lambda_e f(w))}{(\delta(w)+\lambda_r+\lambda_e(1-F(w)))^2}\frac{L}{w} \tag{D.9}$$

$$= \frac{(1-u)g(w)\lambda_e(\delta(w)+\lambda_r+\lambda_e(1-F(w))) - (\lambda_u u + (\lambda_r+\lambda_e G(w))(1-u))(\delta'(w)-\lambda_e f(w))}{\delta(w)+\lambda_r+\lambda_e(1-F(w))}\frac{w}{\lambda_u u + (\lambda_r+\lambda_e G(w))(1-u)} \tag{D.10}$$

$$= w\frac{(1-u)g(w)\lambda_e(\delta(w)+\lambda_r+\lambda_e(1-F(w))) - (\lambda_u u + (\lambda_r+\lambda_e G(w))(1-u))(\delta'(w)-\lambda_e f(w))}{(\delta(w)+\lambda_r+\lambda_e(1-F(w)))\,(\lambda_u u + (\lambda_r+\lambda_e G(w))(1-u))} \tag{D.11}$$

$$= w\frac{(1-u)g(w)\lambda_e(\delta(w)+\lambda_r+\lambda_e(1-F(w)))}{(\delta(w)+\lambda_r+\lambda_e(1-F(w)))\,(\lambda_u u + (\lambda_r+\lambda_e G(w))(1-u))} - w\frac{(\lambda_u u + (\lambda_r+\lambda_e G(w))(1-u))(\delta'(w)-\lambda_e f(w))}{(\delta(w)+\lambda_r+\lambda_e(1-F(w)))\,(\lambda_u u + (\lambda_r+\lambda_e G(w))(1-u))} \tag{D.12}$$

$$= w\left(\frac{\lambda_e(1-u)g(w)}{\lambda_u u + (\lambda_r+\lambda_e G(w))(1-u)} - \frac{(\delta'(w)-\lambda_e f(w))}{(\delta(w)+\lambda_r+\lambda_e(1-F(w)))}\right) \tag{D.13}$$

$$= w\left(\frac{\lambda_e(1-u)g(w)}{\lambda_u u + (\lambda_r+\lambda_e G(w))(1-u)} + \frac{\lambda_e f(w)}{(\delta(w)+\lambda_r+\lambda_e(1-F(w)))} - \frac{\delta'(w)}{(\delta(w)+\lambda_r+\lambda_e(1-F(w)))}\right) \tag{D.14}$$

D.1 Flows from Manning

Manning's firm size is

$$L(w) = \frac{\delta\lambda}{\left(\delta+\lambda(1-F(w))\right)^2}$$

Hence, we'll do the derivative if δ doesn't depend on w and if it does and then we'll figure out which way F has to go such that the direction of bias is amplified and/or offset.

$$\begin{aligned}\frac{\partial L(w)}{\partial w} &= \frac{\delta\lambda 2((\delta+\lambda(1-F(w))))\lambda f(w)}{(\delta+\lambda(1-F(w)))^4} \\ &= \frac{\delta\lambda 2\lambda f(w)}{(\delta+\lambda(1-F(w)))^3}\end{aligned}$$

So the elasticity is

$$\begin{aligned}\varepsilon_L(w) &= \frac{\delta\lambda 2\lambda f(w)}{(\delta+\lambda(1-F(w)))^3}\frac{w}{N(w)} \\ &= \frac{\delta\lambda 2\lambda f(w)}{(\delta+\lambda(1-F(w)))^3}\frac{w}{\frac{\delta\lambda}{(\delta+\lambda(1-F(w)))^2}} \\ &= \frac{\delta\lambda 2\lambda f(w)}{\delta+\lambda(1-F(w))}\frac{w}{\delta\lambda} \\ &= \frac{2\lambda f(w)w}{\delta+\lambda(1-F(w))}\end{aligned}$$

Now supposing that $\frac{\partial\delta(w)}{\partial w} \neq 0$

$$\begin{aligned}\frac{\partial L(w)}{\partial w} &= \frac{\delta'(w)\lambda\left[\delta(w)+\lambda(1-F(w))\right]^2 - \delta(w)\lambda 2\left[\delta(w)+\lambda(1-F(w))\right]\left(\delta'(w)-\lambda f(w)\right)}{\left[\delta(w)+\lambda(1-F(w))\right]^4} \\ &= \frac{\delta'(w)\lambda\left[\delta+\lambda(1-F(w))\right] - \delta(w)\lambda 2\left(\delta'(w)-\lambda f(w)\right)}{\left[\delta(w)+\lambda(1-F(w))\right]^3}\end{aligned}$$

So now the elasticity

$$\begin{aligned}
\varepsilon_L(w) &= \frac{\delta'(w)\lambda\left[\delta(w)+\lambda(1-F(w))\right]-\delta(w)\lambda 2\left(\delta'(w)-\lambda f(w)\right)}{\left[\delta(w)+\lambda(1-F(w))\right]^3}\frac{w}{N(w)} \\
&= \frac{\delta'(w)\lambda\left[\delta(w)+\lambda(1-F(w))\right]-\delta(w)\lambda 2\left(\delta'(w)-\lambda f(w)\right)}{\left[\delta(w)+\lambda(1-F(w))\right]^3}\frac{w}{\frac{\delta(w)\lambda}{(\delta(w)+\lambda(1-F(w)))^2}} \\
&= \frac{\delta'(w)\left[\delta(w)+\lambda(1-F(w))\right]-2\delta(w)\left(\delta'(w)-\lambda f(w)\right)}{\delta(w)+\lambda(1-F(w))}\frac{w}{\delta(w)} \\
&= \frac{\delta'(w)\delta(w)+\delta'(w)\lambda(1-F(w))-2\delta(w)\delta'(w)+2\delta(w)\lambda f(w)}{\delta(w)+\lambda(1-F(w))}\frac{w}{\delta(w)} \\
&= \frac{\delta'(w)\lambda(1-F(w))-\delta(w)\delta'(w)+2\delta(w)\lambda f(w)}{\delta(w)+\lambda(1-F(w))}\frac{w}{\delta(w)} \\
&= \frac{\delta'(w)\lambda(1-F(w))-\delta(w)\delta'(w)}{\delta(w)+\lambda(1-F(w))}\frac{w}{\delta(w)}+\frac{2\lambda f(w)w}{\delta(w)+\lambda(1-F(w))} \\
&= \frac{\lambda(1-F(w))-\delta(w)}{\delta(w)+\lambda(1-F(w))}\frac{w\delta'(w)}{\delta(w)}+\frac{2\lambda f(w)w}{\delta(w)+\lambda(1-F(w))}
\end{aligned}$$

Rearranging this last term:

$$\varepsilon_L(w) = \frac{w\delta'(w)}{\delta(w)}\frac{\lambda(1-F(w))-\delta(w)}{\lambda(1-F(w))+\delta(w)}+\frac{2\lambda f(w)w}{\delta(w)+\lambda(1-F(w))}$$

So the second term is our elasticity with respect to wage in the original model. For the first term, $\frac{w\delta'(w)}{\delta(w)}$ is an elasticity of separations to wages, that we know to be negative. For very high w, $(1-F(w))$ is nearly 0, and so the next term becomes -1, and so it's increasing the elasticity. If $\lambda < \delta(\underline{w})$, the δ of the lowest wage firm, then this is still a negative number.

E Monte Carlo evidence on empirically decomposing separation rate heterogeneity

Endogenous sorting places workers with high separation rates (δ) into low-wage firms. To analyze this, we can estimate empirical separation rates $d_{i,t}$ using a log-linear model that includes worker and firm fixed effects.

$$\log \hat{d}_{i,t} = \gamma_i + \phi_{j(i,t)} + \upsilon_{i,t} \ .$$

This model is correctly specified since the true separation rate is log-additively separable. The error term ($\upsilon_{i,t}$) represents randomness from a homogenous Poisson process.

Monte Carlo simulations show that sorting inflates the estimated variance of firm effects on separations $\phi_{j(i,t)}$ by 1.644 times, with the estimated variance at 4.49% compared to the true variance of 2.73%. This highlights how sorting biases our understanding of the role firms play in worker separations.