Employer Growth and Worker Reallocation

Sunghun Cho*

Stony Brook University

Latest version

Abstract

This paper presents novel evidence from a matched employer-employee data set in the US, investigating a close connection between initial worker flow and dispersion of firm size over time. I find that fast-growing firms replace many workers early on while firms with slow growth do not, leading to a persistent and widening size gap between them. These empirical facts imply that the cleansing process at the firm level is central to understanding the dispersion of firm size. To fully account for the tight link between worker and job flow, I introduce a model in which an employer learns about their employees' productivity, replaces unproductive workers, and grows over the life cycle. Fast-growing firms not only hire more workers but have a better composition in the cross-sectional relationship. With this mechanism, the size of the gap is dictated by the persistence of firms' learning ability. This model explains why a firm's dynamism is affected by workers, and highlights the importance of labor market flexibility in firm growth.

JEL Classification Codes: E24, J63, M13, M51

Keywords: Employer Growth, Matched Employer-Employee Data Set, Worker Reallocation, Firm Dynamics

*Email: sunghun.cho.1@stonybrook.edu; Mailing address: Department of Economics, Stony Brook University; 100 Nicolls Rd, Stony Brook, NY 11794. I am deeply indebted to my academic advisors, David Wiczer, Juan-Carlos Conesa, and Alexis Anagnostopoulos. I thank Gabriel Mihalache, Enghin Atalay, Julieta Caunedo, Katarína Borovičková, Hyunjae Kang, Yurim Lee, Seungki Lee, Joonseok Oh, Alessandra Peter, Junghum Park, Sang-Ha Yoon, and Yoon-Gyu Yoon for their valuable comments and help. I also appreciate all participants in the Macroeconomics and Applied Economics Workshop in the department of economics at Stony Brook University, and 96th Western Economic Association International. Any opinions and conclusions expressed herein are those of the author(s) and do not represent the views of the U.S. Census Bureau. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 1819. All results have been reviewed to ensure that no confidential information is disclosed. All errors are mine.

1 Introduction

Why does firm size differ? In "up-or-out" dynamics (Haltiwanger *et al.*, 2018), high productivity firms grow fast while low productivity firms shrink or even exit.[1] The low productivity firms show higher worker turnover because they separate many workers via poaching to other firms or employees quitting (Bilal *et al.*, 2019). The growth potential of these firms, as an outcome of losing productive workers, is lower than other highly productive firms.

This paper finds the opposite pattern in the data in that fast-growing firms show high turnover rates when young. Using matched employer-employee data observations from the Quarterly Workforce Indicators (QWI), I trace out firm size over 10 years after their entry categorizing them by their worker churn when they are young. The main findings show that fast churners grow quickly while slow churners cannot. As firms age, the size gap among them widens and the initial level of reallocation influences their future size unequally. Looking at the distribution of fast and slow churners across firm size also confirms that entry size does not substantially alter the distribution between churners. These results are not consistent with conventional hypotheses relating firm-level productivity and worker turnover. Rather the composition of workers must be changing so that high worker churn leads to growth. To the best of my knowledge, no other papers have established the tight connection between worker replacement and firm size at the employer level.[2]

This paper emphasizes the importance of net change and the composition of worker flows in firm growth. According to Nagypál (2008), 49% of job leavers are job-to-job transitions while the rest go to non-employment. If a firm loses their productive workers through job-to-job transitions, they are more likely to stay small or exit rather than grow in the long term.[3] This job ladder theory implies that higher worker turnover of this sort negatively impacts firm growth, which contradicts the pattern I show in the data. In this paper, I focus on the flows into non-employment where the cleansing process at the employer level is particularly important for firm growth. Finding unproductive workers early on increases total worker turnover, and separating them allows for a higher ratio of productive workers. The change of this composition improves match

[1] Pioneered by Hopenhayn and Rogerson (1993), many studies incorporate this mechanism to explain firm growth.

[2] As in Burgess *et al.* (2000, 2001), replacing workers, or so-called *worker churn*, may take an important role in firm growth and survival because it affects match quality between employers and employees. For instance, an employer that replaces many employees may lose firm-specific capital early on whereas they can improve match quality by cleansing unproductive workers (Burgess *et al.*, 2000, 2001). Thus, worker reallocation when firms are young no longer guarantees higher growth later on.

[3] On the other hand, on average the workers who move to other firms are more productive than those separated to non-employment.

quality and drives firm size and employment growth.

I build a new model in which the gross and net composition of worker flows from multi-worker firms drives firm growth over the life cycle. The learning process is a key component in firm growth. A firm starts with two different labor forces, known and unknown workers. An employer initially does not know whether their unknown workers are productive or not before they start producing. During production, only the productive workers in the unknown group contribute to the total production. Then, the employer learns about the productivity of a fraction of unknown employees if they are "lucky" (*learning shock*); otherwise it retains the rest of unknown workers. The learning shock determines the intensity of layoffs and posting vacancies while the now-revealed productive employees join the pool of known workers.

Gains and losses from employer learning make a firm simultaneously decide to 1) separate now-known unproductive workers, 2) post vacancies to replace them, and 3) promote the productive ones. Separating these unproductive workers is essential for firm growth. This reallocation between the two worker groups explains how a firm grows in this economy: fast learners grow quicker in size than slow learners and subsequently they become "gazelle" firms.[4]

I calibrate the baseline model to the US economy and simulate firms over the same length of time as in the data set. From the cross-sectional outcomes between gross worker flows and employment growth, I find that this economy has a similar relationship as reported in Davis *et al.* (2013). Expanding firms have a higher hiring rate than separation while contracting firms show the opposite. Additionally, the gross separation rate shows the composition change of known and unknown workers over the firm growth. Firms shrink due to the higher separation of known workers than unknowns. The separation ratio declines as firms grow and the pattern is reversed at higher employment growth firms. In doing so, fast-growing firms have a better composition of workers than other firms.

This paper finds that the higher persistence of a learning shock explains the size of the gap between fast and slow churners. With zero persistence, all firms end up with the same size, and thus show no gap between them. As persistence rises, the faster churners grow more than the slower churners and a substantial gap still remains after 10 years. This implies the process of churning induced by firm-level learning is a crucial component in explaining firm growth both theoretically and empirically.

The rest of the introduction continues to the literature review. Section 2 explains the sample selection in the data set and defines replacement hire. Section 3 introduces the main empirical findings. Section 4 outlines the theoretical framework that depicts how

[4] I borrow this term from Sterk *et al.* (2021), denoting a small share of the high-growth potential of young firms contributing the aggregate job creation later on.

multi-worker firms learn and grow. Section 5 addresses the calibration strategy for the baseline model, and section 6 reports the cross-sectional and longitudinal outcomes in firm simulation. Section 7 concludes.

Literature This paper contributes to the understanding of the nature of potential entrants and their post-entry behavior in studying a firm's dynamism. Pioneered by Hopenhayn and Rogerson (1993), many studies try to explain the dispersion of firm size. Sterk *et al.* (2021) incorporate ex-ante shocks in the standard firm dynamics model and Choi *et al.* (2019) investigate how the characteristics of business founders affect firm growth later on. This paper emphasizes the post-entry behavior of young firms as employers grow by learning their employees' productivity. Sedláček (2020) also studies post-entry behavior varied with aggregate conditions, but he focuses on cohort effects taking the aggregated fluctuation exogenous. In this paper, I find that fast-growing firms have higher worker churn when young and the fast churners grow quicker than other firms after pooling all cohort firms. Thus, I introduce the mechanism in which a firm endogenously reallocates its workers to account for the disparity of firm size between all firms.

A large body of empirical literature investigates a relationship between worker and job flows since Davis and Haltiwanger (1995), Lane *et al.* (1996), and Burgess *et al.* (2000, 2001). Tanaka *et al.* (2020) find a V-shaped pattern of excess hiring, as the same measure in this paper, over the firm employment growth using matched observations sampled from the QWI.[5] Bachmann *et al.* (2021) confirm a similar pattern for the German economy. Such results imply that expanding or contracting firms actively replace their workers. This paper traces out firm size in the long term and categorizes them by worker churn when they young using the panel structure. I find the unequal speed of reallocation leads to a persistent and widening gap between firms. This implies that the distribution of firm size substantially differs by worker churn early on in the V-shaped pattern in Tanaka *et al.* (2020) or Bachmann *et al.* (2021).

This paper also contributes to the understanding of job and worker flow in the multi-worker firm model. Previous studies explain the relationship between job and worker flow by marginal surplus sharing between a firm and multiple workers. (see Bilal *et al.* (2019), Elsby *et al.* (2020), Elsby and Gottfries (2019), Audoly (2019), Gavazza *et al.* (2018), Schaal (2017), Baydur (2017), Fujita and Nakajima (2016), Kaas and Kircher (2015), Acemoglu and Hawkins (2014), Elsby and Michaels (2013), and Cahuc *et al.* (2008)). In doing so, excess hiring is not an optimal choice in their framework. Ex-

[5] Burgess *et al.* (2000) also find a V-shaped pattern of worker churn over job flow, but they use a slightly different measure for worker churn. Their measure includes all excess hiring and separation flow whille this paper only considers excess hiring flow.

cess hiring flow is costly and time-consuming as in a canonical model of Acemoglu and Hawkins (2014). Elsby *et al.* (2020) and Bilal *et al.* (2019) allow firms to have worker churn, but fast-growing firms do not have any incentive to reallocate workers because their composition of workers does not affect firm size later on. Job and worker flow in their model do not allow excess hiring. This paper incorporates a heterogeneous type of workers when young firms start their production. Excess hiring occurs due to the different learning abilities between employers as in Jovanovic (1979) and Nagypál (2007),[6] and fast learners grow quicker than other firms. The mechanism in this paper explains how young firms grow by learning and subsequently such firms become "gazelle" firms in the economy.

2 Data

The firm-level and quarterly observations in the QWI are calculated from the matched employer-employee data set in the Longitudinal Employer-Household Dynamics (LEHD). In this data set, all employers are uniquely identified by State Employer Identification Number (SEIN), and a researcher can see their job and worker flow and other characteristics in the panel structure. The QWI only covers the private sector, but over 95% of jobs are collected in the database (Abowd *et al.*, 2006). This paper uses observations across 17 states in the US from 1990 to 2014.[7]

2.1 Sample Selection

Research samples are selected via the following criteria. First, small firms, which hire fewer than 10 employees, are excluded from the sample.[8] Next, I compute establishment age from the calendar quarters when the establishment starts their business with

[6]Nagypál (2007) distinguishes two types of learning associated with job tenure, "learning by doing" and "learning about match quality." The learning process in this paper incorporates these two features. Match quality reveals out after the production, but an employer only separates those unproductive workers as their learning ability allows. This learning process combines the discrete realization of "learning about match quality" and the error component from "learning by doing" in Nagypál (2007). Thus, I use the term "learning by doing" to indicate the interaction between the two learning processes.

[7]I can only use observations in states approved for the project by Census Bureau. The list of states is: "California, Colorado, Hawaii, Idaho, Illinois, Indiana, Kansas, Maine, Maryland, Missouri, Montana, Nevada, North Dakota, Tennessee, Texas, Virginia, and Washington." (Tanaka *et al.*, 2020)

[8]It is well-known fact that the distribution of employment size has the highest peak around zero, and many establishments stay the same size over their life cycle (Elsby and Michaels, 2013). These employers also present higher worker flow than others because of their small size on average. For instance, suppose an employer hires two workers and separates one worker between period t and $t-1$. If their average employment size was 100, then the hiring and separation rate would be 2% and 1%, but if their average size was 2, then the hiring and separation rate would be 100% and 50%, respectively. To avoid this, I drop out of *too small* employers as in Tanaka *et al.* (2020).

a non-negative number of employees, and drop out the establishment observations if they are initiated by older firms.[9]

I additionally restrict the sample observations that are observed at least once in a calendar year. This allows an establishment to exit for a while in a current year, but they show up in the next calendar year.[10] That is, the length of observed quarters varies from 11 to 44 quarters.[11] Outlier observations in the top 1 percent of separation and hiring rates are deleted.

2.2 Key Variables

The excess accession rate (EAR) measures excess hiring flow beyond the job gains.[12] Elsby *et al.* (2020) and Acharya and Wee (2020) use a similar variable, and it is only a "half side" of worker churn in Burgess *et al.* (2001) (see also Tanaka *et al.* (2020)). Thus, the EAR is computed as follows:

$$EAR_{i,t} = \frac{H_{i,t} - \max\left(H_{i,t} - S_{i,t}, 0\right)}{\frac{1}{2}\left(E_{i,t} + E_{i,t-4}\right)} \tag{1}$$

where $H_{i,t}$, $S_{i,t}$ and $E_{i,t}$ denote the number of hires, separations and employed workers respectively in period t by firm i. The denominator follows the standard normalization as in Davis *et al.* (1998). If an employer in t hires employees greater than their job creation, then they separate the same number of workers in the same period. In doing so, the greater extent of the EAR implies that the employer "replaces" many workers, not fully captured in the job change. If the gross hire and separation flow are equal to the job creation, the EAR is zero.[13] This variable computes the probability of employees being replaced in their current employer rather than identifying "who has been replaced by whom." The EAR in the main findings is averaged out over the first three

[9] That is to say, establishments activated by state-level firm age greater than 2 are not included in the sample. From now on, I interchangeably use the terms firm and establishment.

[10] Theoretically, a firm can exit (or just not recorded in the data set) at most three quarters. For instance, if this firm has been recorded in Q1 at the calendar year t and does not have an employment record in Q2, Q3, and Q4, then they must show up in the next calendar year $t+1$.

[11] Although the selection of this sort allows a short exit of firms, it is not completely free of impact from the permanent exit. In the data, many firms often exit (either permanently or shortly) and young and small firms are more prone to exit over their first three years (Sterk *et al.*, 2021). Therefore, it may under- or overestimate the result.

[12] Accessions in this term imply that all new hires at the current quarter, no matter how long they have been employed thus far or employed again this quarter, "recalls" (Tanaka *et al.*, 2020)

[13] The EAR does not imply that employee A in period t has been replaced by a new employee B in the same period. Firms may hire workers ahead of the termination, on time, or later. Thus, this measure does not consider the sequence or timing of hire and separation. Kuhn and Yu (2021) list four types of replacement hire: "Early Refills, On-time Refills, Temporary Replacements and Late Refills," based on employee's "departure time." The EAR includes all these types of replacement hires which can be overlapped between periods.

years of firm age to track the initial average of worker churn.

The speed of worker reallocation depends on the distribution of firms for each categorical group of firm age and worker churn. For instance, 30% of the EAR can be fast enough for the top tenth percentile in firms aged over 10 years while it is more likely to be the median level in younger firms. Therefore, the terms "fast" or "slow" churners indicate the relative intensity of their replacement *within* a group.

This approach has a couple of limitations. Firms may hire or replace their workers at different timing. That is, a firm may replace their worker ahead of job creation or hire workers later due to the delayed schedule. I reduce this bias to take the average of excess hiring over the first three years after firms enter. Furthermore, I check the rank of individual churners when they are young and older, and their rank does not substantially change. Young firms usually have higher worker churn and it declines as they age as in Burgess *et al.* (2001).

Worker churn does not distinguish the case either 1) the employer loses productive workers (by demand shock or workers are poached) or 2) the employer separates unproductive workers.[14] Other types of unobserved heterogeneity may exist. For future work, it is important to investigate the employee's characteristics and decompose the aggregated numbers.

3 Empirical Findings

Figure 1 presents novel evidence between worker churn and firm size, as guessed by Burgess *et al.* (2000) and Burgess *et al.* (2001).[15] The centered gray line represents the average size of all firms. The marginal change in firm size declines as firms age, consistent with a stylized fact in firm growth (Sterk *et al.*, 2021).

I categorize firms by different percentiles of worker churn when their age is 0 to 2, and trace out the annual average of employment size over time. The average size is denominated by firm's initial size to look at the relative growing pattern.[16] The blue line represents that a firm has worker churn in the bottom tenth of overall distribution when young. Slow churners grow only 30% after 10 years of their entry. As the rank of worker churn becomes higher, firms grow quicker in size. The firms in the red line grow more than 60% to their initial size and they have a top tenth of worker churn when

[14]The measure of this sort, unless it is disaggregated further, has a similar issue. To directly observe the replacement between workers, I need a data set containing all hiring schedules of the incumbent and new workers as in Kuhn and Yu (2021).

[15]As mentioned earlier, they predict worker churn may positively affect the firm growth when firms separate their unproductive worker and improve match quality. This cleansing process outweighs the loss of firm-specific capital, so figure 1 shows this pattern.

[16]Weighting the average by employment size does not change the overall trend. See the appendix for additional figures and size distribution before denomination.

they are young.

The gap between fast and slow churners widens and looks persistent. If worker churn is short-term friction, the size gap between firms is more likely to disappear in the long term as in Acemoglu and Hawkins (2014) The gap of the size in figure 1 contradicts the prediction in previous studies. Rather it shows a diverging pattern between firms.[17]

Cohort and industry-fixed effects may affect this result. Sedláček (2020) shows how young entrants during the Great Recession have less potential to create jobs than other entrants in the normal period. They quantitatively show that the aggregate condition drives the persistent differences between cohorts. In figure 1, all firms are pooled out so that this cancels off the impact from particular cohorts during the economic downturn. Barrero *et al.* (2020) also find that many firms, during the COVID-19 pandemic, reallocate their workers in the business survey. This implies that worker churn at the firm level substantially differs regardless of entry timing.

The relationship between worker churn and firm size may be changed in different industries. Kuhn and Yu (2021) estimate the cost of worker turnover in the retail industry is lower than in other industries. In the QWI, the gross rate of worker churn differs by industry, implying the ex-ante heterogeneity of the potential entrants. For instance, an employer's replacement cost is higher in the finance or information sectors. Their entry decision and post reallocation can be affected by the heterogeneity of this sort.[18]

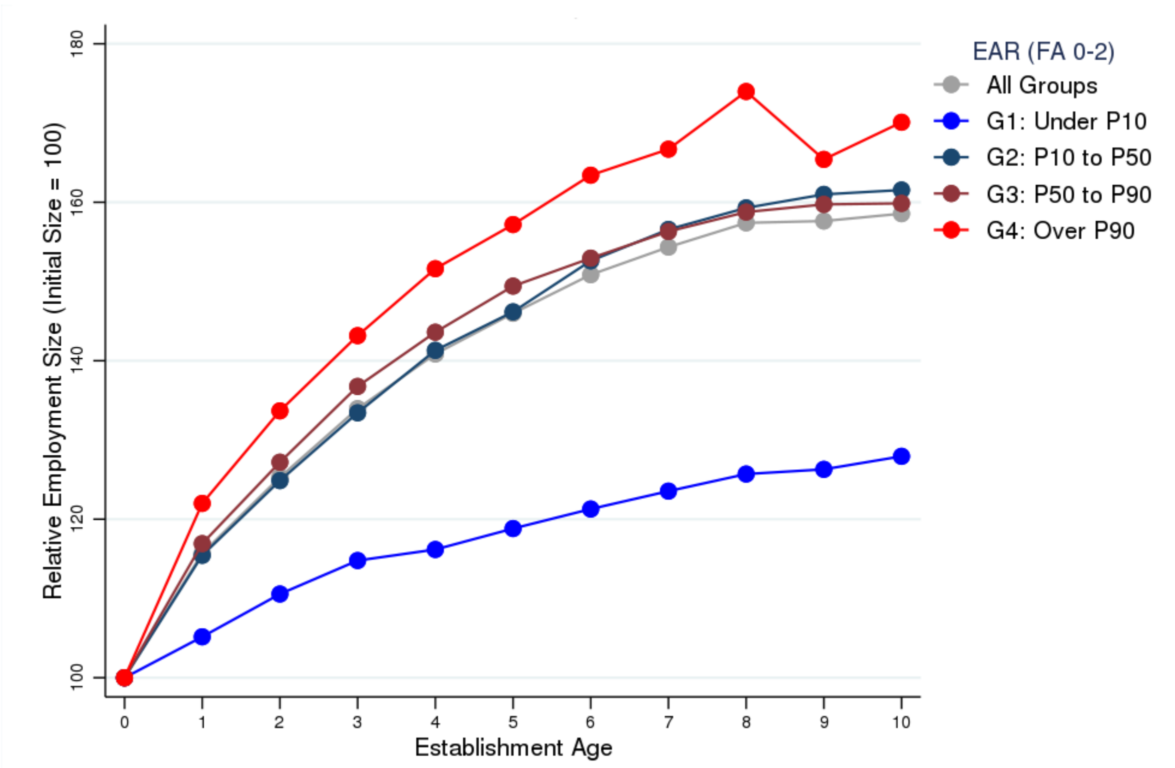
The entry size of firms may substantially alter the rank of worker churn. I compute the percentile thresholds of worker churn by categorizing the firm's entry size in figure 2. Overall, the entry size does not substantially change the percentile threshold of fast and slow churners. The average threshold of 85 to 90 percentile is greater than 40% for all sizes. Similarly, all the averages of 5 to 10 percentile are lower than 5%. The average threshold of the median group, 45 to 50 percentile, declines more than 10% in 250-499 and 500+ size bin. The distribution of worker churn when firms are young is slightly right-skewed.

To sum up, fast-growing firms have a higher rate of worker churn when they are young. This also leads to a persistent size difference in the long run. The empirical results in the two figures contradict the conventional hypotheses relating to productivity and worker churn at the firm level (Hopenhayn and Rogerson, 1993; Acemoglu and

[17] According to Acemoglu and Hawkins (2014), ex-post heterogeneity, such as adjustment cost or demand shock, makes firm size differ from its optimal target. Firms grow unequally, but they end up the same size in the long run. Since this figure absents the ex-ante differences, the diverging pattern can be a "snapshot" of transitioning a firm's long-run size. As in Sterk *et al.*, 2021, the heterogeneity before entry is more likely to create a permanent difference. But worker churn herein is a post-entry decision of young firms, so I will introduce a model incorporating the mechanism of entrant's growth affected by worker flow.

[18] I plot residual coefficients of age dummies regressed on the EAR after taking out the industry and cohort effect. I confirm that the residual plot shows a similar pattern in figure 1.

Figure 1: Trajectory of Firm Size by Initial Replacement Hires



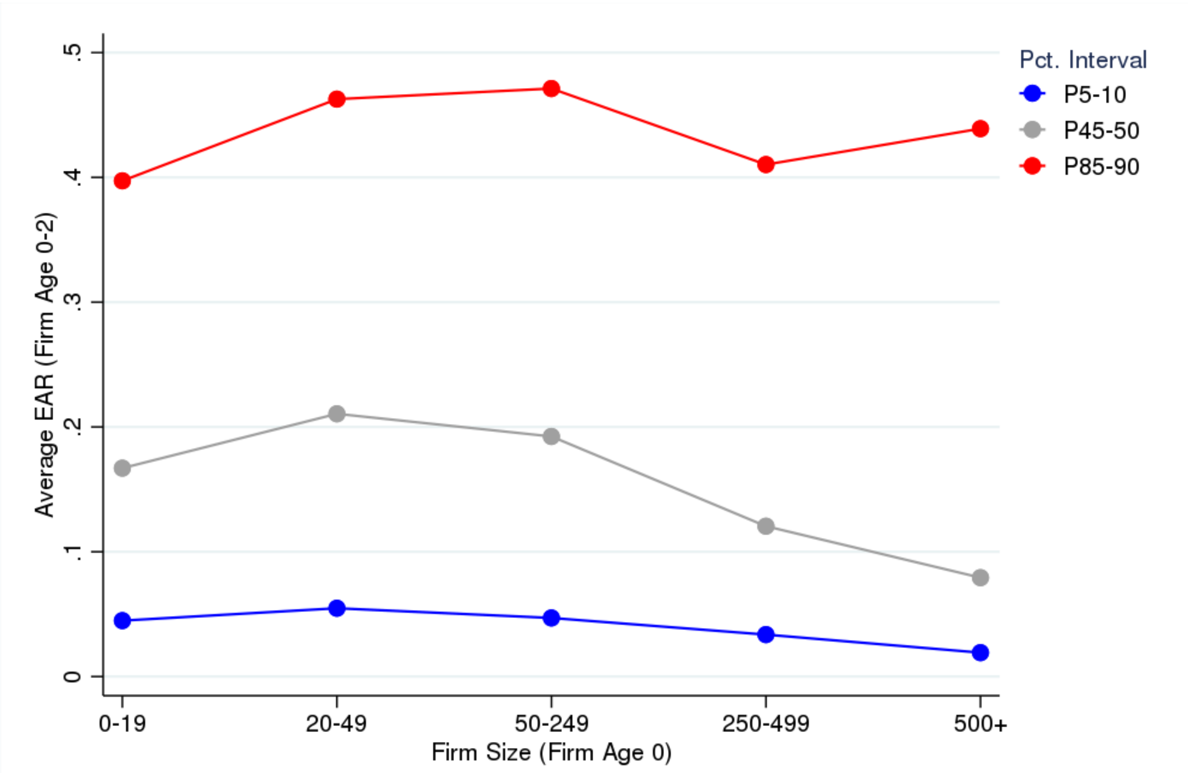
Note: Each dot represents the unweighted average of firm size for each group of replacement hires and firm age in all industries. I denominate the firm size by its average size in the first year after entry. The gray line is the average employment size of all firms. These firms are categorized by worker churn when their age is 0 to 2. Firm's initial worker churn is less than 10 percentile in the blue line (G1), greater than 10 percentile and below the median in the navy line (G2), greater than the median and below 90 percentile in the maroon line (G3), and greater than 90 percentile in the red line (G4).

Hawkins, 2014; Bilal *et al.*, 2019; Elsby *et al.*, 2020). Rather, such results imply that the worker composition must be changing as firms grow. Employers replace many workers early on to improve match quality, leading to high worker churn and subsequent employment growth. I will develop a theoretical framework in which a firm changes its worker composition based on its learning ability in the next section.

4 Theoretical Model

I build a multi-worker firm model in which employers have different learning abilities in the frictional labor market. A firm in this model produces one single output with

Figure 2: Initial Distribution of Replacement Hires by Firm Size



Note: Each dot represents the weighted average of replacement hire when firm age is 0 to 2 for categorical sizes in all industries. This average is a *pseudo* number of percentile intervals to avoid disclosing confidential information. I weight the average by firm size when its age is zero. The categorical firm size follows the standard definition in the Business Dynamics Statistics (BDS).

two different types of labor forces: known and unknown workers. While known workers are fully productive, the productivity of unknown workers has not been revealed out before production starts. The employer learns their unknown employee's productivity after they observe the total production level, so this is "learning-by-doing." The employer separates unproductive workers and posts vacancies while the now-revealed productive workers are promoted. Fast learners have a better composition of workers and grow quicker in size than slow learners.

I will show that the model mechanism matches the key empirical moments in the data. In the simulation, the firm's transitional path is dictated by the different abilities of learning. The persistence of learning shock accounts for the size of the gap between firms in figure 1. Starting with the overall model environment, I will describe a firm's value problem with wage determination, and aggregate variables.

4.1 Model Overview

In this economy, a firm starts with two types of employed workers, known workers (n) and unknown workers (n_0). When an employer hires a new worker from non-employment, they do not know the productivity level of those workers. The match quality thus is unknown to employers before they observe the level of total production. At the beginning of period t, the firm receives an idiosyncratic shock x and new firms enter. The mass of entrants will be normalized to one as in Restuccia and Rogerson (2008).

The employer receives a productivity signal $(y \in \{\overline{y}, \underline{y}\})$ from unknown workers after production. The signaling probability $(Pr(y = \overline{y}) + Pr(y = \underline{y}) = 1)$ pins down the number of workers who send the signal. However, the employer only learns a fraction of them because their learning ability is disproportionately limited by the learning shock (ξ). The higher shock of learning they receive, the more unknown workers they will learn. They separate unproductive workers and promote productive workers to a new part of known workers.[19]

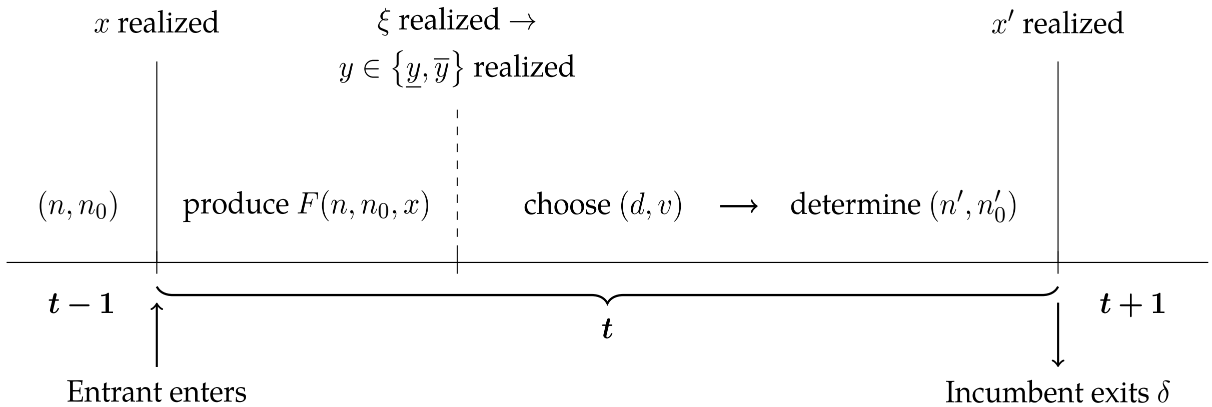
Firms post vacancies (v) after separating unproductive workers. They replace those workers to fill out the size gap due to the layoff of unknown workers. This vacancy posting creates a replacement hiring flow. Meanwhile, they may also separate more workers after learning because it is expensive to retain all labor forces. They choose additional separation (d) between known and unknown workers stayed after learning shock realized. In doing so, employers decide to hire and separate workers at the same moment. They consider both the total size and the composition of workers in this mechanism. Firms grow quicker in size when they have a better composition. Lastly, the incumbent firms exit at a constant rate (δ) before the next period comes in. Figure 3 outlines the infinite sequence of a firm's decision and learning.

4.2 Matching Environment

The matching environment follows the standard setup. There are many firms and workers that randomly meet in the labor market by matching technology. A mass of incumbent firms is μ, and a mass of entrants is m_e that is normalized to one. Firms compete for hiring workers from unemployed while employed workers are terminated either by learning or they quit at a constant rate (s). The total size of the labor force is $\mathcal{L} = E + \mathcal{U}$

[19] Silva and Toledo (2009) also introduce two labor forces, entrant and incumbent workers in their model, and firms know the fraction of entrant worker's productivity at a constant rate. This paper, in contrast, incorporates a different learning ability between firms, so the fraction of workers they learn is changing over time.

Figure 3: Model Timeline



Note: t denotes the time period. I use the prime notation to indicate the next period in the model. n is the number of known workers and n_0 is the number of unknown workers. x is an idiosyncratic shock in which a firm receives at the beginning of the period t. $F(\cdot)$ is a production function. ξ is a learning shock which determines the intensity of the employer's learning ability. y is a true signal of employees' productivity whether it turns out to be a low productivity, $\underline{y}$, or a high productivity, $\overline{y}$. d is a policy for separating workers and v is a policy for posting vacancies. δ is a constant exit rate same for all firms.

where E is the total number of employed workers and $\mathcal{U}$ the total number of unemployed workers. Employed workers cannot search on the job until they are separated. The aggregate vacancy posted is $\mathcal{V}$. The new match Z follows the matching technology with constant returns to scale.

$$Z = Z(\mathcal{U}, \mathcal{V}) \tag{2}$$

The market tightness is $\theta \equiv \frac{\mathcal{V}}{\mathcal{U}}$. Then, the vacancy filling rate is $q(\theta) \equiv Z\left(\frac{1}{\theta}, 1\right)$ (a vacancy meets a worker) and the job finding rate is $p(\theta) \equiv Z(1, \theta)$ (a worker meets a vacancy).

4.3 Firm's Value

Firm-level learning is an important component in this model. The realization of learning shock ξ in period t changes their flow profit and discounted future values. The forward-looking firms, paired with their idiosyncratic shocks and costs of input and adjustment, decide their optimal policies.

A firm produces their single goods with known and unknown workers and then pay the cost of labor inputs equally. The wage only compensates their employees' unemployment benefits, which will be explained in the next section. Firms pay various type of costs when they adjust their size by posting vacancies or separating workers.[20] These two policies determine the number of workers employed in the next period, and pin down the flow profit.

The path of unknown workers, who become finally known to employers, is twofold. First, unknown workers who become unproductive are separated immediately from their current employer and go back to the unemployed. Second, unknown workers who become productive are promoted to a part of known workers. The rest of the unknown workers stay for the next period unless they are additionally separated.

Firms determine the number of vacancies posted to hire better workers from outside. Meanwhile, this fills the gap of the layoff in the production. In the next period, the newly hired workers are added to the unknown part. Both known and unknown workers quit their employer at a constant rate s, so it is "worker-initiated" quit. d_1 and d_2 are "employer-initiated" quit (Silva and Toledo, 2009), leading to endogenous separation flow. Therefore, the full description of the firm's problem is

$$\begin{aligned} V\left(n, n_0, x, \xi\right) = \max_{v, d_1, d_2} \Bigg\{ & \underbrace{F\left(n, n_0, x\right)}_{\text{Production}} \\ & - \underbrace{w\left(x\right)\left(n + n_0\right)}_{\text{Labor Cost}} - \underbrace{C_v\left(v\right)}_{\text{Vacancy Posting Cost}} - \underbrace{C_d\left(n, n_0, d_1, d_2\right)}_{\text{Separation Cost}} \\ & + \beta\left(1-\delta\right) E_{x', \xi' | x, \xi} V\left(n', n_0', x', \xi'\right) \Bigg\} \end{aligned} \tag{3}$$

subject to

$$n' = \left(1 - d_1\right) n + \xi n_0 \cdot Pr\big(y = \overline{y}\big) \tag{4}$$

$$n_0' = \left(1 - d_2\right)\left(1 - \xi\right) n_0 + v q\left(\theta\right) \tag{5}$$

where $Pr\big(y = \overline{y}\big)$ is a probability of productive type, $w\left(\cdot\right)$ is a flow wage, δ is an exogenous exit rate, and $q\left(\theta\right)$ is a vacancy filling rate. The vacancy policy v is bounded below zero, $v \geq 0$, and the separation policy is bounded as $s \leq d_1, d_2 \leq 1$.

The cost structure of posting a vacancy is convex, as in Kaas and Kircher (2015) and Bilal *et al.* (2019), to pin down the distribution of firm size in equilibrium. A firm also pays the extra cost when they separate workers more than the exogenous outflow. This is a similar way to define firing cost as in Mukoyama and Osotimehin (2019), Elsby *et al.* (2020), and Hopenhayn and Rogerson (1993) because a firm can avoid the separation cost if they choose $d_1 = s$, $d_2 = s$, or both.

[20] Employers jump to their optimal size unless they have a ex-post friction as such. This paper incorporates different learning ability between firms, making a persistent disparity in size in the long term.

4.4 Determination of Worker's Wage

The firm in this model evenly pays their known and unknown workers. This wage compensates the unemployment benefit as a base, and pays an extra amount of money depending on the realization of their idiosyncratic shock.

$$w(x) = b(1+x) \tag{6}$$

This wage rate is a special case of wage scheme from the marginal surplus sharing as in Elsby and Michaels (2013); Elsby and Gottfries (2019); Elsby *et al.* (2020); Stole and Zwiebel (1996).[21] I use this setup for a couple of benefits. In traditional Nash wage bargaining, a firm competes to hire marginal workers and the wage rate depends on the bargaining power between workers and a firm. But the bargaining in two groups of workers and a firm is a non-trivial problem. As raised in Stole and Zwiebel (1996) and Brügemann *et al.* (2019), infra-marginal wage in a large size firm comes with a complex sequence of bargaining. Taking the simplest setup for wage, this paper avoids the complication of this sort. I can also assume away worker's value problem for this setup because the monopsonistic firm has all bargaining power.

4.5 Firm Distribution

As in Hopenhayn and Rogerson (1993) and Restuccia and Rogerson (2008), this paper track the firm distribution as follows:

$$\mu(n', n_0', x', \xi') = (1-\delta) P' \mu(n, n_0, x, \xi) + m_e \cdot \pi_e(n, n_0, x, \xi) \tag{7}$$

where μ is the mass of incumbent firms, m_e is the mass of entrants, P is a transition matrix implied from the firm's optimal policy and π_e is an invariant distribution of entrants. If m_e is normalized to one, then the entry rate of entrants is scaled to the exogenous exit of incumbent firms.

[21] Pioneered by Stole and Zwiebel (1996), Elsby and Gottfries (2019) derive the following wage scheme from the marginal surplus sharing.

$$w = \frac{\beta}{1-\beta(1-\alpha)} x \alpha n^{\alpha-1} + \omega_0$$

where $\omega_0 \equiv \beta\omega_f + (1-\beta)\omega_e$. If $\alpha = \beta = 1$ (i.e., a firm has all bargaining power), then

$$w = x + \omega_f$$

where ω_f is a benefit for a firm when the negotiation between current workers and their employer breaks down. If I put $\omega_f = b(1+x) - x$, then this is equivalent to the equation 6.

4.6 Aggregation

Given the stationary measure $\mu^*(n, n_0, x, \xi)$, the total number of vacancies is added up all posting vacancies of individual firms.

$$\mathcal{V}^* = \int g_v(n, n_0, x, \xi)\, d\mu^*(n, n_0, x, \xi) \tag{8}$$

As in Elsby and Michaels (2013), the job creation and Beveridge curve determine the number of unemployed workers in equilibrium. In the Beveridge curve, there are two parts of unemployed workers. One is those who are separated from the employer either by learning or endogenous separation and the other is unemployed workers who cannot find a job in this period. Those two workers enter the next period as a new mass of total unemployed.

$$\mathcal{U}' = (1 - p(\theta))\mathcal{U} + \int \left\{ g_{d_1}(n, n_0, x, \xi)\, n + \big((1-\xi)\, g_{d_2}(n, n_0, x, \xi) + \xi Pr\big(y = \underline{y}\big)\big) n_0 \right\} d\mu(n, n_0, x, \xi)$$

Then, in a steady state, the increase in job finding rate decreases the total number of unemployed while the workers currently separated increases the total.

$$\mathcal{U}^* = \frac{1}{p(\theta)} \underbrace{\int \left\{ g_{d_1}(n, n_0, x, \xi)\, n + \big((1-\xi)\, g_{d_2}(n, n_0, x, \xi) + \xi Pr\big(y = \underline{y}\big)\big) n_0 \right\} d\mu^*(n, n_0, x, \xi)}_{\text{Total Separation}} \tag{9}$$

From the definition of the total labor force, the residual amount of taking out total hires implies another number of unemployed workers in this economy.

$$\mathcal{U}^* = \mathcal{L} - \underbrace{\int \left\{ g_n(n, n_0, x, \xi) + g_{n_0}(n, n_0, x, \xi) \right\} d\mu^*(n, n_0, x, \xi)}_{\text{Total Hire}} \tag{10}$$

The market tightness in stationary equilibrium is pinned down by these two relationships.[22]

5 Calibration

I calibrate the model to the US economy. The model frequency is monthly to replicate high volatility of worker flow (Elsby and Michaels, 2013; Elsby *et al.*, 2020). Some parameters are directly assigned by the data set or existing literature while the rest of the

[22] Instead of solving general equilibrium, this paper takes a partial equilibrium approach. Given the empirical value of the market tightness and job finding rate, I match all other moments to get a reasonable value of vacancy filling rate or so.

parameters is internally calibrated to match the empirical target. A set of moments in the model are aggregated from monthly to quarterly observations to match or compare the moments only quarterly provided in the LEHD.

Parameters Fixed from Data Set or Model Assumption The matching function is a Cobb-Douglas form with constant returns to scale as in Elsby and Michaels (2013) and Bilal *et al.* (2019).

$$Z = \epsilon_1 \mathcal{U}^{\epsilon_2} \mathcal{V}^{1-\epsilon_2} \tag{11}$$

In doing so, the vacancy filling rate is $q(\theta) \equiv \epsilon_1 \theta^{-\epsilon_2}$ and the job finding rate is $p(\theta) \equiv \epsilon_1 \theta^{1-\epsilon_2}$. In Petrongolo and Pissarides (2001), the matching elasticity is close to 0.5 in the data, so I take this value to ϵ_2. I also assign $\theta = 0.72$ (Pissarides, 2009) and $p(\theta) = 0.4$ (Bilal *et al.*, 2019) from literature to match the model to the data set. Putting them altogether, I calculate $\epsilon_1 = \frac{p(\theta)}{\theta^{1-\epsilon_2}} \fallingdotseq 0.4714$, and subsequently, the vacancy filling rate is $q(\theta) \fallingdotseq 0.5556$.

The discount rate, β, is chosen to match the annual 5% interest rate in the US. From the monthly interest rate 0.4%, the discount rate is thus $\beta = \frac{1}{1+0.004} \approx 0.996$.

I choose the Constant Elasticity of Substitution (CES) production function for two reasons. First, two labor forces in the model are complements if the substitution elasticity is less than one. Firms now have an incentive to employ both known and unknown workers for their production. Otherwise, an employer would be better off choosing a corner solution. For instance, they can separate all unknown workers and retain only known workers. Thus, this ensures the interior solution of worker composition. However, the complementarity does not bound the total size.[23] I add the decreasing returns to scale to pin down their total size in equilibrium.

[23] Suppose there is no adjustment cost and a firm only pays labor input cost. Then, the marginal value of increasing the number of n or n_0 workers is

$$\frac{\partial\left(\left(\omega n^{\frac{\sigma-1}{\sigma}} + (1-\omega) n_0^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} - w(n+n_0)\right)}{\partial n} : \left(\omega n^{\frac{\sigma-1}{\sigma}} + (1-\omega) n_0^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}-1} \omega n^{\frac{\sigma}{\sigma-1}-1} - w = 0$$

$$\frac{\partial\left(\left(\omega n^{\frac{\sigma-1}{\sigma}} + (1-\omega) n_0^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} - w(n+n_0)\right)}{\partial n_0} : \left(\omega n^{\frac{\sigma-1}{\sigma}} + (1-\omega) n_0^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}-1} (1-\omega) n_0^{\frac{\sigma}{\sigma-1}-1} - w = 0$$

Since workers are paid equally, then these two equations will be rearranged as follows.

$$\therefore \frac{n_0}{n} = \left(\frac{\omega}{1-\omega}\right)^{-\sigma}$$

The weight and substitution elasticity determine the optimal path of worker composition as a firm grows. The dispersion of the composition depends on the degree of adjustment costs. But this does not limit the total size because a firm can grow infinitely until the resource allows.

More importantly, unknown workers contribute to the total production disproportionately before an employer learns their true productivity. If unknown and unproductive workers also increase the production, their employer cannot distinguish where this increment came from. For this reason, I assume only "good" workers change the production. That is to say, the productive type of unknown workers actually increases the total production.[24]

$$F\left(n, n_0, x\right) = x\left(\overline{y}\right)^{\alpha}\left(\omega n^{\rho} + \left(1-\omega\right)\left(Pr\left(y=\overline{y}\right)n_0\right)\right)^{\frac{\alpha}{\rho}} \tag{12}$$

where $\rho = \frac{\sigma-1}{\sigma}$, $0 < \sigma < 1$ and $0 < \omega < 1$.[25]

The exogenous exit rate is chosen to match the monthly exit rate in the BDS. From 1990 to 2014, the annual exit rate on average, out of the total number of establishments, is close to 10%. Then, the monthly exit rate δ is 0.83%, divided by 12 months.

A constant fraction of workers are separated from their current employer in every period. This exogenous outflow is set to match the quit share of the monthly EU rate, 1.6%, reported in Elsby *et al.* (2020) and Bilal *et al.* (2019). Since 22.5% of workers voluntarily quit from their employer out of total quit, the worker outflow s is $0.225 \times 0.016 = 0.0036$.

I normalize the unemployment benefit as suggested in Elsby and Gottfries (2019). In the Mortensen-Pissarides search model, the reservation wage level equals the discounted value of unemployment benefit. As such, the reservation level is $w^* = \frac{b}{1-\beta(1-\delta)}$ in this model. If $w^* = 1$ as in Elsby and Gottfries (2019), then $b = 1 - \beta\left(1-\delta\right) = 1 - 0.996 \times (1 - 0.0083) = 0.012267$.

I take the same convex form of posting vacancies as in Borovickova (2016).

$$C_v\left(v\right) = \frac{c_1}{1+c_2}\left(v\right)^{1+c_2} \tag{13}$$

where c_1 is a scale parameter and c_2 denotes a shape parameter. I fix $c_2 = 2$ for a quadratic shape in marginal vacancy posting and internally calibrate c_1.

I set the cost function for separating workers as follows:

$$C_d\left(n, n_0, d_1, d_2\right) = \begin{cases} \tau \cdot \left\{\left(d_1 - s\right)n + \left(d_2 - s\right)n_0\right\} & \text{if } s < d_1, d_2 \leq 1 \\ 0 & \text{if } d_1, d_2 = s \end{cases} \tag{14}$$

where τ is a common parameter for both types of separation. A firm pays extra separation cost only if they separated workers greater than the constant outflow.

[24] See the appendix for omitted steps to derive the equation 12.

[25] I set the CES form and boundary conditions for σ and ω mainly based on Acemoglu and Restrepo (2020) and Dolado *et al.* (2021).

In Silva and Toledo (2009), the firing cost in the US is estimated to be the 8 weeks of wage payment according to the World Bank data source introduced in their paper. To match this, τ is chosen to be the twice average monthly wage in the model, $2 \times Ew = 2 \times b\,(1+1) = 0.0491$.

The idiosyncratic shock x follows the standard AR(1) form as in Elsby *et al.* (2020) and Baydur (2017), then

$$\ln x' = \rho_x \ln x + \varepsilon'_x \tag{15}$$

where $\varepsilon'_x \sim N\left(0, \sigma_x^2\right)$. Sterk *et al.* (2021) estimate a set of parameters to replicate ex-post shocks at the establishment level using LEHD. I adopt their estimates for the persistence of this shock, $\rho_x = 0.963$, but choose σ_x^2 to prevent negative flow profits across all the states in the firm's value.

The learning shock, ξ, controls the intensity of employer learning, but none of the data set has such information on the employer side in the US.[26] I assume the learning shock follows AR(1) process as in the idiosyncratic shock and matching the persistence and variance gives rise to the same grid of learning intensity in a different model simulation.[27]

$$\ln \xi' = \rho_\xi \ln \xi + \varepsilon'_\xi \tag{16}$$

where $\varepsilon'_\xi \sim N\left(0, \sigma_\xi^2\right)$. As the grid comes in for the simulation, different rates (from 5% to 16%) of unknown workers are separated from their current employer by learning. Out of total separation, the fraction of separation by firm-level learning explains 40% to 50% on the average of all firms in different degrees of the persistence. In quantitative exercises, I will show that the different rate of persistence accounts for the substantial gap between firms using the same grid of learning shock.

Internally Calibrated Parameters The rest of the parameters are internally calibrated to match their empirical targets. Since all parameters are jointly moved, I only explain the closest target for each parameter. Two parameters in the CES production, as shown before, determine the slope of the optimal path in frictionless setup. The weight param-

[26] There are a couple of data sources broadly related. According to the Mass Layoff Statistics program in U.S. Bureau of Labor Statistics (this survey was terminated after 2013), the mass layoff in private non-farm or manufacturing sectors accounts for 15% to 22% out of total separation in one quarter (Bureau of Labor Statistics, 2012). This data set provides the number of separations by categorical reasons, but none of them is directly related to the learning intensity. In fact, learning may exist in all those categories. Another source is the Displacement Worker Survey, reported every three years for supplements of the Current Population Survey. This also offers specific reasons for the displacement of workers, but has the same issue in the Mass Layoff Statistics. For instance, the establishment closing (41%), insufficient work (37%) and position abolished (13%) account for 91% of total displacement in all industries (Kandilov and Kandilov, 2010). All those could have displaced workers initiated by employer learning.

[27] The joint pair is $\left(\rho_\xi, \sigma_\xi\right) \in \left\{(0, 0.15), (0.5, 0.13), (0.75, 0.1), (0.9, 0.06)\right\}$ and all these imply the same grid of learning shock, which is $\xi \in \left\{0.2, 0.35, 0.5, 0.65, 0.8\right\}$.

eter, ω, mainly affects the ratio in aggregate labor. I match this aggregate ratio of the two workers to the share of single to multi-quarter jobs in Hyatt and Spletzer (2017).[28] They report that the average rate of single-quarter jobs is 8% and the multi-quarter jobs are 12.5% from 1996 to 2012 in the LEHD. Then, the aggregate ratio targets 0.64 ($= \frac{8\%}{12.5\%}$).

The labor elasticity jointly matches the correlation between the log employment in the current and previous periods. Sterk *et al.* (2021) recently estimate the covariance structure in firm size over firm age. I calculate the average of those numbers, except for the auto-correlation in the diagonal, and match the substitution elasticity to it. Their average correlation between log employment is 0.852, which is highly persistent.[29]

The hazard rate of unproductive workers are closely related to the size of learning shock and the probability of unproductive types. Given the size grid of learning shock, I match this probability to the hazard rate, implied from the cumulative distribution, within three months.[30] If the firing event occurs at random time T, then $Pr\left(T < t\right) = F\left(t\right)$. Let $S\left(t\right) = 1 - F\left(t\right)$ be the survival function and $h\left(t\right) = \frac{f(t)}{S(t)}$ be the hazard function where $f\left(t\right)$ is a density function of $F\left(t\right)$. If $F\left(t\right) = 1 - \exp\left(-\lambda t\right)$, then the hazard function is $h\left(t\right) = \lambda$. I estimate λ from the average of inflow rate of unknown workers in the simulation.

$$\hat{\lambda} = \frac{1}{N \cdot T} \sum_{i,t} \frac{S_{i,t}}{E_{i,t}} \cdot \mathcal{I}_{\{n_0 = 1\}} \tag{17}$$

where $S_{i,t}$ is a separation flow and $E_{i,t}$ is a employment level in a firm i at time t. In doing so, the hazard rate of unproductive workers in a quarter is computed by

$$F\left(3\right) = 1 - \exp\left(-3\hat{\lambda}\right) \tag{18}$$

This number in the model is now matched to the incidence of single-quarter jobs in Hyatt and Spletzer (2017), which is approximately one-third out of the total hire and separation flow.

Lastly, the scale parameter of vacancy cost, c_1, matches the job creation rate as in Borovickova (2016). In the BDS, the average job creation rate from 1990 to 2014 is 15.3%.

[28] Hyatt and Spletzer (2017) define the single-quarter jobs as "those jobs that begin and end within the same calendar quarter" and the multi-quarter jobs are "jobs last more than one quarter." I assume that firms are more likely to separate unproductive workers by the short spell jobs rather than the longer spell jobs. In the model, these two counterparts are the total number of unknown and known workers, respectively.

[29] Krusell *et al.* (2000) calibrate these parameters with aggregate income share and the wage premium between skilled and unskilled workers. However, I cannot follow their strategy because of the wage scheme in this paper. Two different workers are paid evenly so that the implied income share in aggregates is too small to match the number in the national account. Also, no wage premium exists. For these reasons, I calibrate them with different targets as explained.

[30] I follow the notation from Steven Stern's lecture note.

Table 1: Calibrated Parameters and Empirical Targets

Parameter		**Values**	**Reasons**
Fixed from data or model assumption			
β	Discount rate	0.996	Annual interest rate 5%
α	Returns to scale	0.640	Elsby *et al.* (2020)
$\overline{y}$	Production scale	1.000	Normalization
ϵ_1	Matching efficiency	0.471	Job finding rate = 0.4
ϵ_2	Matching elasticity	0.500	Petrongolo and Pissarides (2001)
δ	Exogenous exit rate	0.008	Monthly exit rate = 0.83%
s	Exogenous worker outflow	0.004	Quit share of EU rate
b	Unemployment benefit	0.012	Normalization
c_2	Shape parameter	2.000	Quadratic cost
τ	Separation cost	0.049	Firing cost = 8 week of earnings
ρ_x	AR(1) shock persistence	0.963	Sterk *et al.* (2021)
σ_x^2	AR(1) shock variance	0.050	Positive flow profit
ρ_ξ	AR(1) learning persistence		$\xi \in \{0.2, 0.35, 0.5, 0.65, 0.8\}$
σ_ξ	AR(1) learning std.		
Internally calibrated in the model			
σ	Labor elasticity	0.200	Corr $(\log E_t, \log E_{t-1}) = 0.852$
ω	Weight for n workers	0.580	$\frac{\text{Avg. rate of single-quarter jobs}}{\text{Avg. rate of multi-quarter jobs}} = 0.64$
$Pr(y = \underline{y})$	Prob. of unproductive type	0.260	Share of single-quarter jobs in total = $\frac{1}{3}$
c_1	Scale parameter	0.400	Quarterly job creation rate = 3.825%

Thus, the quarterly rate is 3.825% $(= \frac{15.3\%}{4})$. I calculate the average of quarterly job gains in the model to match this target. Table 1 lists up all calibrated parameters and their empirical targets.

6 Model Outcomes

6.1 Targeted and Non-Targeted Moment

Table 2 summarizes the model performance. In the baseline case ($\rho_\xi = 0$), the model fits the data well except for the job creation rate.[31] Quarter-to-quarter, the log level of employment is highly persistent as in the data set. The other case of learning persistence also shows a similar extent of correlation. The ratio of known and unknown workers in the model is well matched to the job-spell moment in the LEHD. The hazard rate in

[31] In Elsby and Michaels (2013) and Elsby *et al.* (2020), they calibrate the model to the share of *inactive* firms in the data to replicate a big spike nearby -2% to 2% growth. In this paper, the learning mechanism encourages a firm to grow more than staying the same size, so fewer inactive firms exist. In doing so, the job creation rate can be greater than the standard setup while I can focus more on the transitional dynamics of *growing* firms.

Table 2: Targeted and Non-Targeted Moment

	Data	Model			
		(1)	(2)	(3)	(4)
		$\rho_\xi = 0.0$	$= 0.50$	$= 0.75$	$= 0.90$
A. Targeted Moment					
Correlation of log employment	0.852	0.839	0.701	0.717	0.820
Ratio of single to multi quarter jobs	0.640	0.568	0.569	0.575	0.583
Hazard rate of single quarter jobs	0.333	0.330	0.329	0.327	0.321
Quarterly job creation rate (%)	3.825	7.107	10.895	12.915	11.756
B. Non-targeted Moment					
Employment growth in G1 (%)	2.658	5.798	5.426	4.342	2.847
Employment growth in G4 (%)	5.445	6.872	7.359	7.823	6.385
Size gap between G1 and G4	0.235	0.068	0.121	0.195	0.232

Note: ρ_ξ is a persistence of AR(1) learning shock.

the model is almost equivalent to the target in the data.

The bottom in the same table displays non-targeted moments from the model and its counterpart in the figure 1. I calculate the annualized growth and the size gap between firms in G1 and G4 and omit the rest of the groups.[32] The annualized growth rate between fast and slow churners implies the 23.5% of size gap. The baseline model has almost no difference in the average growth rate between firms, neither does its size gap. However, the rest of the columns shows that their disparity substantially widens as the persistence level increases. This implies that firm rank in worker churn drives the firm growth as such in the data set and I will present the full transitional dynamics in the section 6.3.

6.2 Simulation Results

I simulate 25,000 firms for 11 years of firm age over the same data periods in the figure 1. The time frequency in the model is monthly to capture the high volatility of worker flow as suggested in Elsby and Michaels (2013). The employment and worker flow variables are aggregated up from monthly to quarterly by the same definition in the QWI.[33]

[32]Firm size of churners in G1 ends up 30% greater than their entry size in their age 10, so the annualized rate of growth is $\left(\frac{130}{100}\right)^{\frac{1}{10}} - 1 \fallingdotseq 0.02658$. Similarly, the annualized growth in G4 is $\left(\frac{170}{100}\right)^{\frac{1}{10}} - 1 \fallingdotseq 0.05445$. The size gap is $1 - \frac{130}{170} \fallingdotseq 1 - 0.7647 = 0.2353$.

[33]For details, monthly hires and separations are added up for each quarter. The quarterly employment takes first monthly values (e.g., January, April, July, and October) because QWI uses "Beginning-of-Quarter Employment: Counts" to measure the firm size.

This model takes an exogenous entry and exit, so it is crucial to pin down the firm size at entry in simulation. In the BDS, the average size of incumbent firms is 22.1 while the average size of entrant firms (i.e., firm age is zero) is 10.7. Motivated by this empirical moment, all firms at age 0 employ half of the average size of incumbent firms.[34] Suppose total size is the sum of known and unknown workers ($N = n + n_0$). Then,

$$N_{\text{FirmAge}=0,t=1} = \frac{1}{2} \times \overline{N}$$

Next, I assume a firm draws a ratio of unknown workers to the total size from the bounded uniform distribution because the accurate number of unknown workers is not observable in the data set.

$$\frac{n_0}{N_{\text{FirmAge}=0,t=1}} \sim \mathcal{U}\,(0.75, 0.95)$$

This ratio determines the number of known and unknown workers.[35] The figure 4 shows that the size distribution of young entrants (dotted lines) is right-skewed than incumbent firms (solid line) and their average size increases as firms age.[36]

Figure 5 introduces a cross-sectional relationship between worker flow and employment growth. The Expanding firms have a higher hiring rate than separation while contracting firms show the opposite. Davis *et al.* (2013) report a similar pattern between hire and separation in the data set.[37] Gross change in worker flow directly implies the net change in the job flow. This graph has more lines for the number of known and unknown workers separated from contracting or expanding firms. For instance, contracting firms separate known workers more than unknowns, implying that they are losing productive workers. On the other hand, fast-growing firms, higher than 30% of employment growth, separate unknown workers more than known employees. Although the gross separation is lower than hire in growing firms, there exists heterogeneity of the worker composition. Therefore, a change in the gross and net composition of worker flow gives rise to a change in employment size. Bilal *et al.* (2019) show a similar illustrative relationship between worker and job flow, but their model does not have this composition change when firms grow.

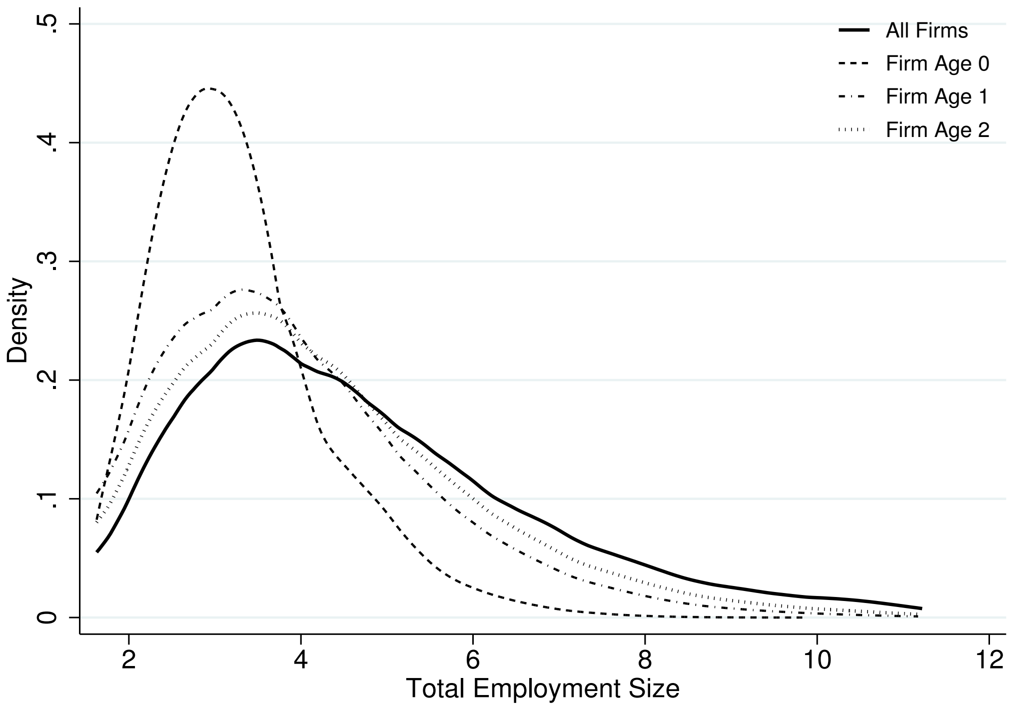
[34] To get the average of incumbent firms, I simulate one firm with many periods before the panel simulation, and take the average of the total size.

[35] The bounded numbers in the uniform distribution are fairly arbitrary, but I also checked that a modest change of those numbers does not significantly alter the results.

[36] In figure 4 and 5, I confirm that the level of learning persistence does not significantly change the overall patterns. Thus, I only introduce the case of $\rho_\xi = 0.75$ for the middle value.

[37] This paper does not have a kinked point as their plot, or so-called "hockey stick." In the model economy, a firm separates many workers by learning and endogenous allocation. Future hires, given the vacancy filling rate, consider all these separations as in a mirror relationship.

Figure 4: Size Distribution of Entrants and Incumbents



Note: This graph displays kernel density estimates using Epanechnikov algorithm when the persistence of employers' learning ability is 0.75. Different persistence does not substantially alter the overall distribution. I select 0.5 for bandwidth of the kernel estimates. The distribution of firm size is right skewed when firm age is zero, and the right tail becomes fatter with a slight increase of the mean as firms age. The solid black line represents the distribution of all firms in the simulation.

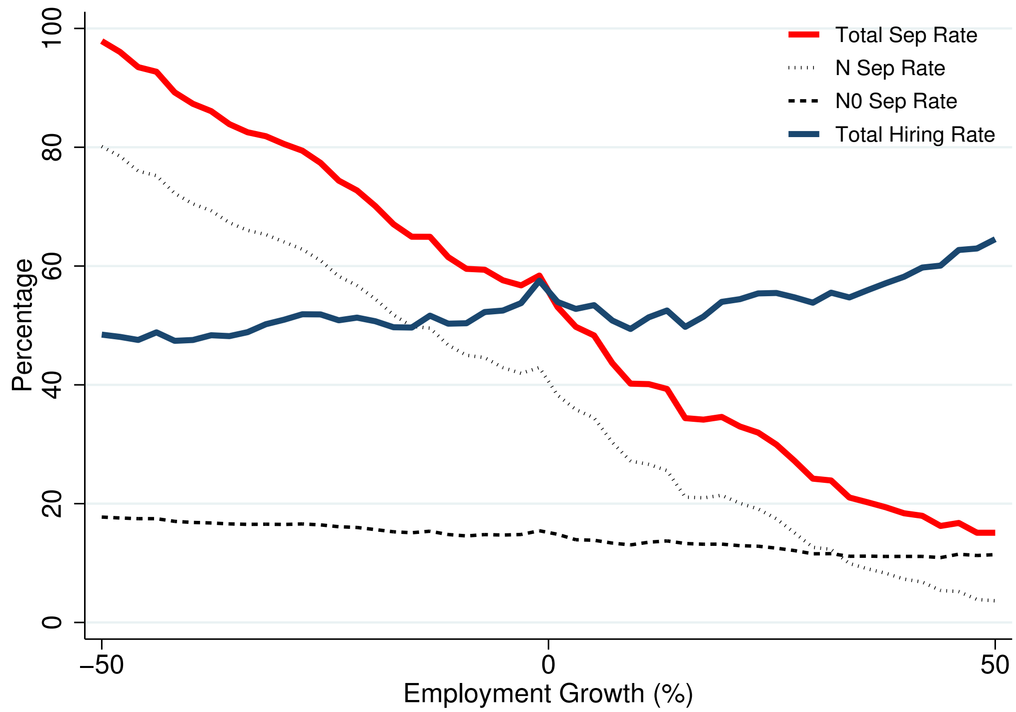
6.3 Source of Persistent Difference

I simulate the baseline model with different persistence of learning shocks to account for the size gap in figure 1. Higher persistence implies that young firms know better how fast they separate unproductive workers in the future. The figure 6 displays the transitional path of fast and slow churners in red and blue lines equivalently defined in the previous figure.

If the firm's learning process is randomly decided, the initial worker churn does not change the size later on. The solid red and blue lines converge to a similar size after 10 years. Non-targeting moments in table 2 support this growth pattern. When a firm has no persistence of learning ability, the size of the gap between fast and slow churns is only 6% while the figure 1 shows 23.5%.

The faster churners grow more than the slow churners and their gap remains after

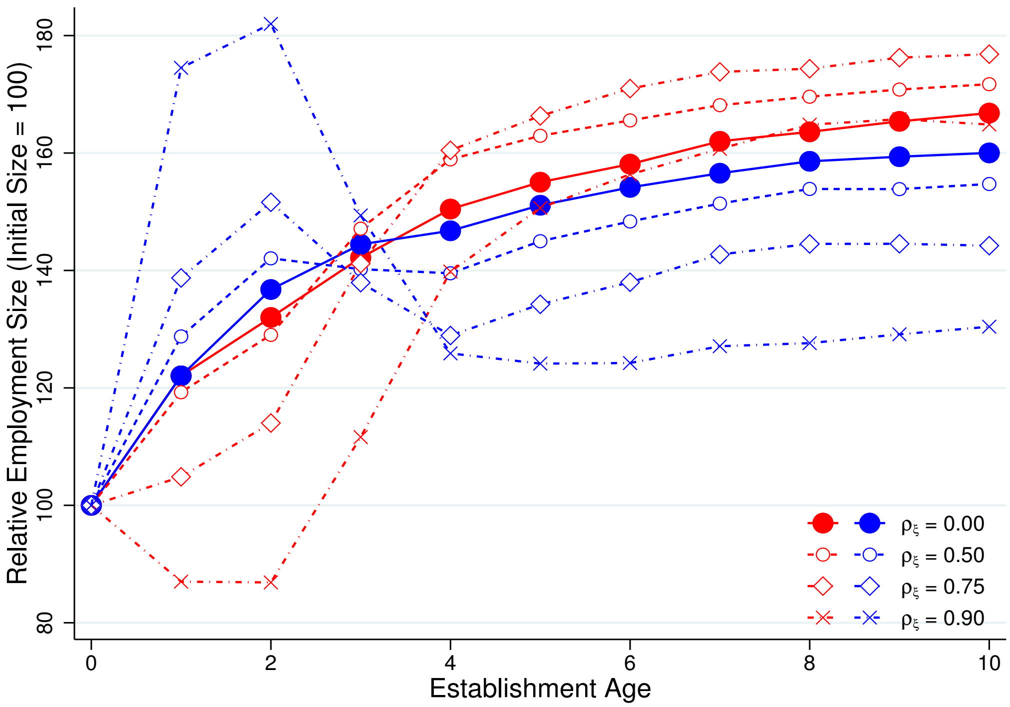
Figure 5: Worker Flow over Employment Growth



Note: This graph displays the cross-sectional relationship between worker flow and employment growth in firm simulation when the persistence of employers' learning ability is 0.75. Different persistence does not substantially alter the overall shape. The range of the horizontal axis is restricted from -50% to 50%. The red line denotes the total separation rate and the navy line is the total hiring rate. The dotted and dashed lines represent the number of known and unknown workers separated over employment growth. The ratio between known and unknown workers declines as firms grow and the pattern is reversed at higher employment growth firms, around 30%.

10 years as the persistence of learning ability rises. According to the table 2, the increasing persistence accounts for over 50-90% of the gap in firm size. Firms also show a different transitional path as their learning ability changes. Slow churners present "overshooting" in firm size during the first four years and steeply declines later on. They have avoided costs of size adjustments due to smaller loss of unknown workers, but unproductive workers are piled up. Then, they cannot produce higher than before and it negatively affects their size in the long run. Fast churners grow slowly or downsize due to the higher adjustment cost of learning. It is unavoidable, but they hold more productive workers and separate unproductive workers. With the better composition of workers, they outgrow the slow churners and stay on the top of the size ladder in the long term.

Figure 6: Persistence of Firms' Learning Ability and the Size Gap between Firms



Note: I plot the top tenth churners in the red line and the bottom tenth churners in the blue line as in figure 1. Each marker is the average firm size for each group of firm age and worker churn. The shape of markers represents the different persistence of firms' learning ability in AR(1) learning shock. As persistence rises, the faster churners grow more than the slower churners and a substantial gap still remains after 10 years.

7 Conclusion

This paper answers how firm size differs by worker flow and what makes the persistent difference over time. Unlike the conventional hypothesis relating to productivity and worker turnover, fast-growing firms show high worker churn when they are young, leading to a persistent size gap in the long term. This paper emphasizes the importance of changing worker composition, and the cleansing process at the firm level to account for this new finding. In the multi-worker firm model, the gross and net composition of worker flows drives firm growth over the life cycle. The employer improves match quality by learning their employees' productivity, which is a key component in the model. Worker reallocation occurs when the employers simultaneously separate unproductive workers and post vacancies for recruiting better workers. This leads fast churners to grow quicker in size than slow churners. In the simulation, I show that expanding firms have a better composition of their workers than shrinking firms. Also, the greater extent of an employer's learning ability accounts for the size of the gap between fast and slow churners in the main empirical findings.

This paper explains why firm growth is affected by worker flows. Furthermore, Davis and Haltiwanger (2014) emphasize the importance of "labor market fluidity." They argue that the declining labor reallocation may lead to lower employment growth in the US economy. Thus, it is more important to restore the fluidity in the labor market for economic growth. As such, this paper highlights the influential impact of worker churn on firm growth, representing the labor market flexibility. The reduced fluidity in the labor market impedes the cleansing process at the employer level, which a firm has a better composition of workers and grows later on. The market fluidity, in this regard, is also closely related to the birth of "gazelle" firms.

For future work, the worker churn can be decomposed by the worker's wage growth or employment status in job spells.[38] This decomposition shows which group of workers has been replaced over the firm's life cycle as in Kuhn and Yu (2021). Young firms are hard to grow if they lose their workers due to them being poached to other firms. The share of job-to-job transitions to total turnover is higher in this case. A set of reasons for worker separation also helps understand whether employers fire those workers due to their low productivity (learning) or lose them for other reasons (poached or quit).

Ex-ante differences of potential entrants, pioneered by Sterk *et al.* (2021), are another interesting direction. Potential entrants in the wholesale industry tend to hire lower-skill workers than in the finance industry. Worker churn, with the ex-ante disparities, may have an uneven effect on firm growth. The model in this paper helps explain why the different nature of entrants affects their post-entry behavior and the distribution of firm size in the long term.

The model can be improved with the following setups. In the general equilibrium, the separated workers immediately affect the market tightness and search intensity from the unemployed status. It is interesting to check the impact of general equilibrium on heterogeneous firms and their size growth.

All workers are not allowed to search on the job in the model. They are also paid evenly regardless of their productivity level. I may consider the wage bargaining protocol as in Elsby and Gottfries (2019), Elsby *et al.* (2020), and Brügemann *et al.* (2019) incorporating a realistic wage setup. They have developed a bargaining wage in a multi-worker firm model allowing for on-the-job search. In doing so, productive workers have

[38] Tanaka *et al.* (2020) combine the sample of job spells at the worker level with a job and worker flow information from the QWI to study workers' earnings growth. This only looks at the change in the worker side. As for the firm side, I need a representative sample of firms in the QWI and match it to the worker characteristics observed in job spell data. Then, the worker-side information is aggregated at the employer level. A researcher now decomposes the gross worker churn into several categorical groups. For instance, we can see the number of workers separated to the non-employment in the next job spell or whether their wage is increased or not.

a better wage scheme than unproductive workers in terms of their outside options and marginal contribution to the total production.

References

Abowd, J., Stevens, B., Vilhuber, L., Andersson, F., McKinney, K., Roemer, M. and Woodock, S. (2006) The lehd infrastructure files and the creation of the quarterly workforce indicators. lehd technical paper no, *Washington, DC: US Census Bureau*.

Acemoglu, D. and Hawkins, W. B. (2014) Search with multi-worker firms, *Theoretical Economics*, **9**, 583–628.

Acemoglu, D. and Restrepo, P. (2020) Unpacking skill bias: Automation and new tasks, in *AEA Papers and Proceedings*, vol. 110, pp. 356–61.

Acharya, S. and Wee, S. L. (2020) On-the-job search and the productivity-wage gap.

Audoly, R. (2019) Firm dynamics and random search over the business cycle.

Bachmann, R., Bayer, C., Merkl, C., Seth, S., Stüber, H. and Wellschmied, F. (2021) Worker churn in the cross section and over time: New evidence from germany, *Journal of Monetary Economics*, **117**, 781–797.

Barrero, J. M., Bloom, N. and Davis, S. J. (2020) Covid-19 is also a reallocation shock.

Baydur, I. (2017) Worker selection, hiring, and vacancies, *American Economic Journal: Macroeconomics*, **9**, 88–127.

Bilal, A. G., Engbom, N., Mongey, S. and Violante, G. L. (2019) Firm and worker dynamics in a frictional labor market.

Borovickova, K. (2016) Job flows, worker flows and labor market policies.

Brügemann, B., Gautier, P. and Menzio, G. (2019) Intra firm bargaining and shapley values, *The Review of Economic Studies*, **86**, 564–592.

Bureau of Labor Statistics, U. D. o. L. (2012) Extended mass layoffs, third quarter 2012, *The Economics Daily*.

Burgess, S., Lane, J. and Stevens, D. (2000) Job flows, worker flows, and churning, *Journal of labor economics*, **18**, 473–502.

Burgess, S., Lane, J. and Stevens, D. (2001) Churning dynamics: an analysis of hires and separations at the employer level, *Labour Economics*, **8**, 1–14.

Cahuc, P., Marque, F. and Wasmer, E. (2008) A theory of wages and labor demand with intra-firm bargaining and matching frictions, *International Economic Review*, **49**, 943–972.

Choi, J., Goldschlag, N., Haltiwanger, J. and Kim, J. D. (2019) The post entry growth and survival of business startups: The role of founding teams, *Unpublished Manuscript*.

Davis, S. J., Faberman, R. J. and Haltiwanger, J. C. (2013) The establishment-level behavior of vacancies and hiring, *The Quarterly Journal of Economics*, **128**, 581–622.

Davis, S. J. and Haltiwanger, J. (1995) Measuring gross worker and job flows.

Davis, S. J. and Haltiwanger, J. (2014) Labor market fluidity and economic performance.

Davis, S. J., Haltiwanger, J. C., Schuh, S. *et al.* (1998) Job creation and destruction, *MIT Press Books*, **1**.

Dolado, J. J., Motyovszki, G. and Pappa, E. (2021) Monetary policy and inequality under labor market frictions and capital-skill complementarity, *American Economic Journal: Macroeconomics*, **13**, 292–332.

Elsby, M. and Gottfries, A. (2019) Firm dynamics, on-the-job search and labor market fluctuations.

Elsby, M., Michaels, R. and Ratner, D. (2020) Vacancy chains.

Elsby, M. W. and Michaels, R. (2013) Marginal jobs, heterogeneous firms, and unemployment flows, *American Economic Journal: Macroeconomics*, **5**, 1–48.

Fujita, S. and Nakajima, M. (2016) Worker flows and job flows: A quantitative investigation, *Review of Economic Dynamics*, **22**, 1–20.

Gavazza, A., Mongey, S. and Violante, G. L. (2018) Aggregate recruiting intensity, *American Economic Review*, **108**, 2088–2127.

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

Hopenhayn, H. and Rogerson, R. (1993) Job turnover and policy evaluation: A general equilibrium analysis, *Journal of political Economy*, **101**, 915–938.

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

Jovanovic, B. (1979) Job matching and the theory of turnover, *Journal of political economy*, **87**, 972–990.

Kaas, L. and Kircher, P. (2015) Efficient firm dynamics in a frictional labor market, *American Economic Review*, **105**, 3030–60.

Kandilov, A. M. and Kandilov, I. T. (2010) Job displacement from agriculture, *American Journal of Agricultural Economics*, **92**, 591–607.

Krusell, P., Ohanian, L. E., Ríos-Rull, J.-V. and Violante, G. L. (2000) Capital-skill complementarity and inequality: A macroeconomic analysis, *Econometrica*, **68**, 1029–1053.

Kuhn, P. and Yu, L. (2021) How costly is turnover? evidence from retail, *Journal of Labor Economics*, **39**, 461–496.

Lane, J., Stevens, D. and Burgess, S. (1996) Worker and job flows, *Economics Letters*, **51**, 109–113.

Mukoyama, T. and Osotimehin, S. (2019) Barriers to reallocation and economic growth: the effects of firing costs, *American Economic Journal: Macroeconomics*, **11**, 235–70.

Nagypál, É. (2007) Learning by doing vs. learning about match quality: Can we tell them apart?, *The Review of Economic Studies*, **74**, 537–566.

Nagypál, É. (2008) Worker reallocation over the business cycle: The importance of employer-to-employer transitions, *Manuscript, Northwestern Univ*.

Petrongolo, B. and Pissarides, C. A. (2001) Looking into the black box: A survey of the matching function, *Journal of Economic literature*, **39**, 390–431.

Pissarides, C. A. (2009) The unemployment volatility puzzle: Is wage stickiness the answer?, *Econometrica*, **77**, 1339–1369.

Restuccia, D. and Rogerson, R. (2008) Policy distortions and aggregate productivity with heterogeneous establishments, *Review of Economic dynamics*, **11**, 707–720.

Schaal, E. (2017) Uncertainty and unemployment, *Econometrica*, **85**, 1675–1721.

Sedláček, P. (2020) Lost generations of firms and aggregate labor market dynamics, *Journal of Monetary Economics*, **111**, 16–31.

Silva, J. I. and Toledo, M. (2009) Labor turnover costs and the cyclical behavior of vacancies and unemployment, *Macroeconomic Dynamics*, **13**, 76–96.

Sterk, V., Sedláček, P. and Pugsley, B. (2021) The nature of firm growth, *American Economic Review*, **111**, 547–79.

Stole, L. A. and Zwiebel, J. (1996) Intra-firm bargaining under non-binding contracts, *The Review of Economic Studies*, **63**, 375–410.

Tanaka, S., Warren, L., Wiczer, D. *et al.* (2020) Earnings growth, job flows and churn.

Appendix

A Additional figures

Figure A.1: Trajectory of Average Firm Size by Initial Replacement Hires (All Industries, Unweighted)

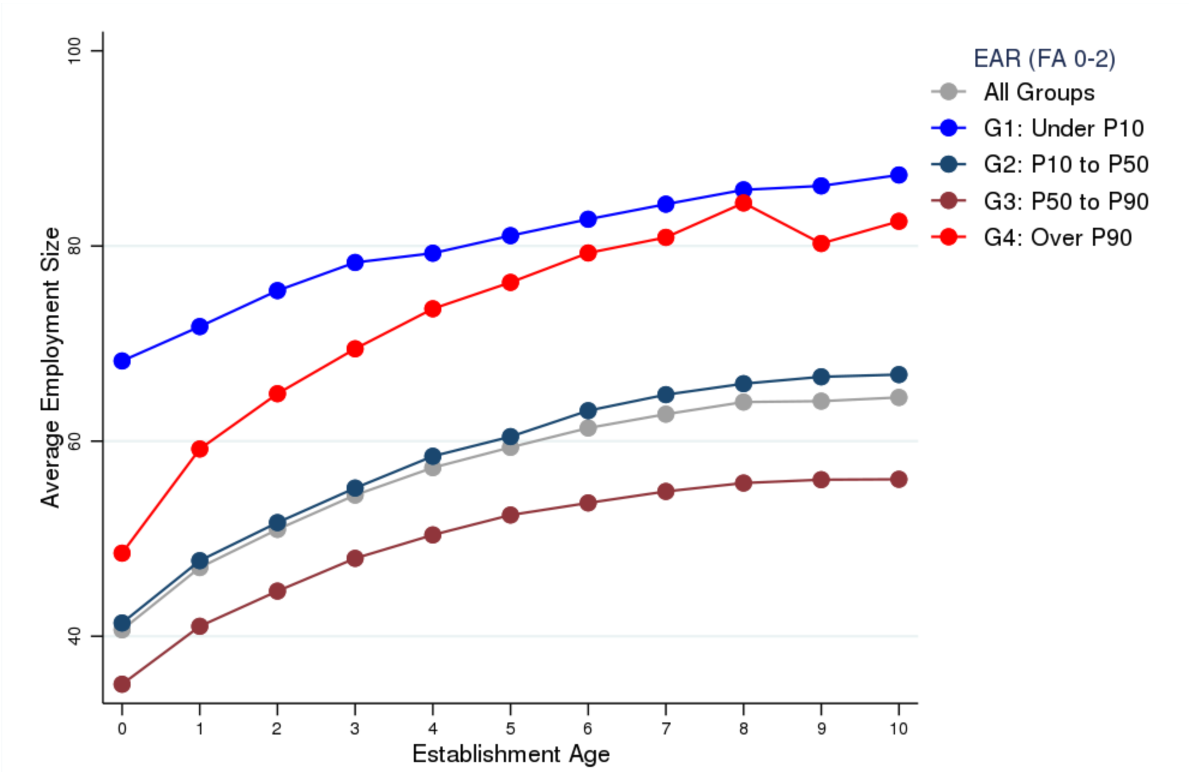


Figure A.2: Trajectory of Average Firm Size by Initial Replacement Hires (All Industries, Weighted)

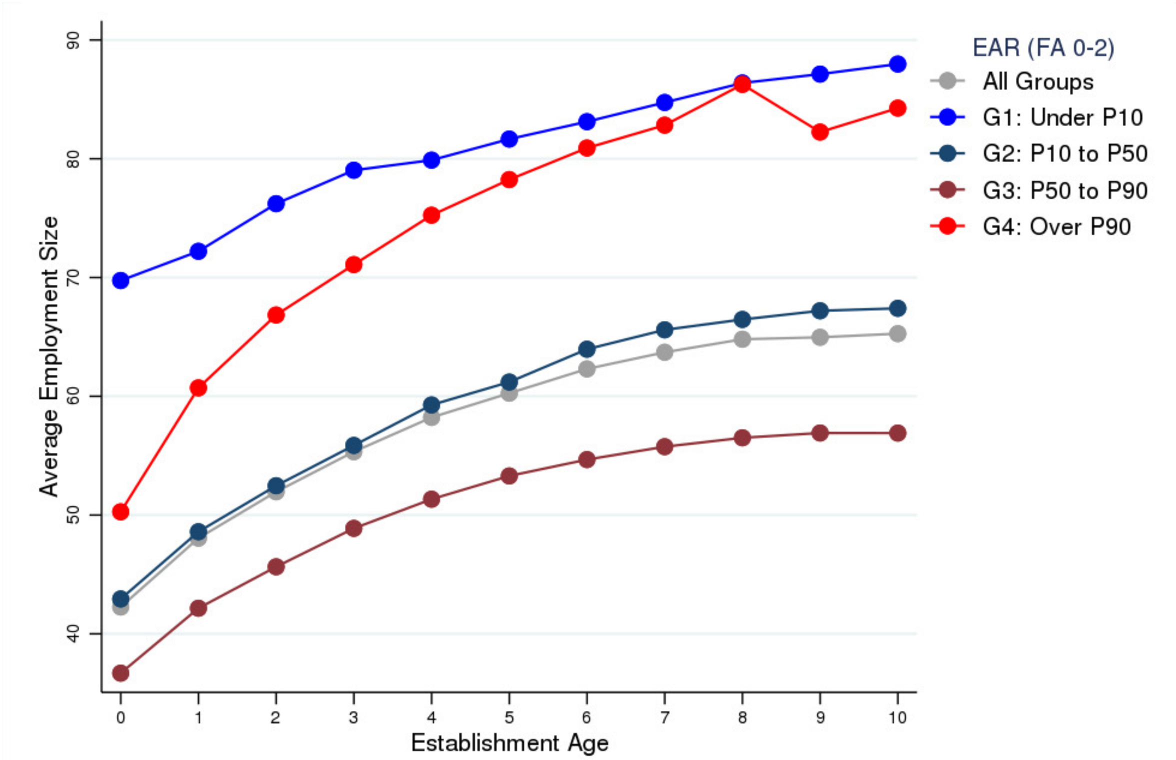
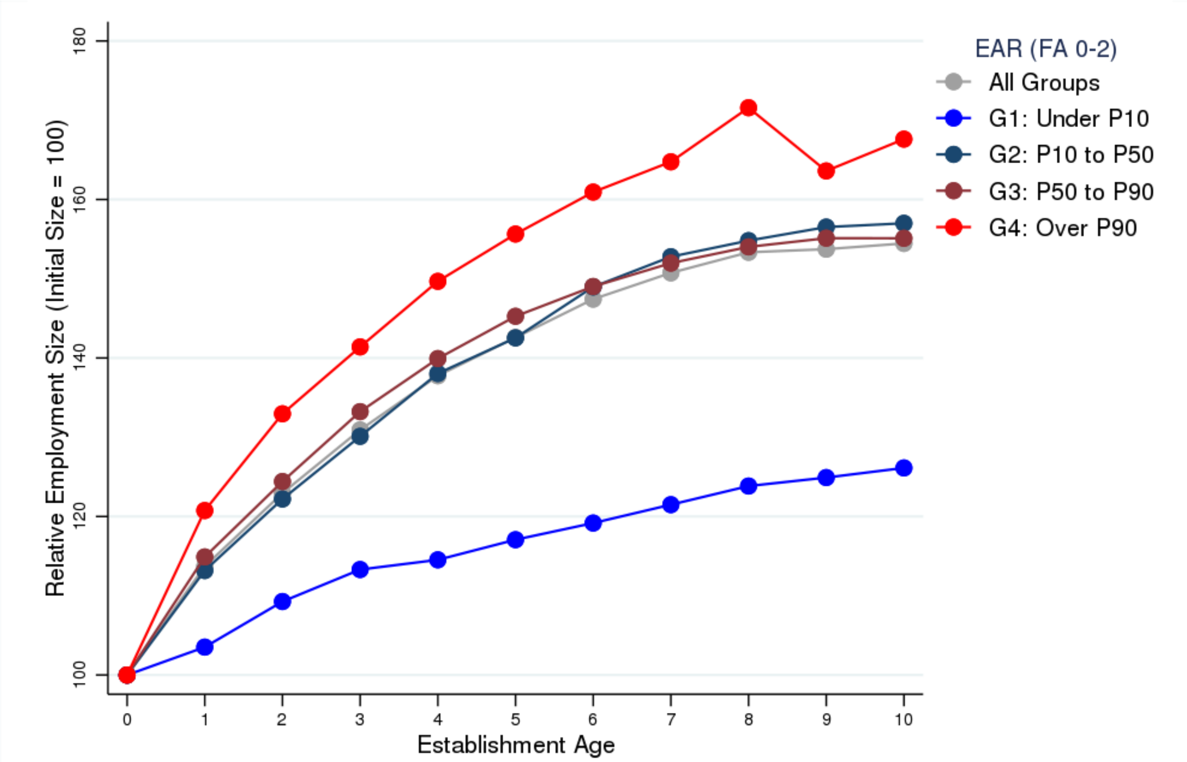


Figure A.3: Trajectory of Relative Firm Size by Initial Replacement Hires (All Industries, Weighted)



B Production Function

Suppose there is no adjustment cost, but only exist labor cost for two labor forces in the firm's value. For simplification, only consider the marginal value of flow profit for each worker.

$$V\left(n, n_0, x\right) = F\left(n, n_0, x\right) - w\left(x\right)\left(n + n_0\right)$$

The cross derivative for n and n_0 only leaves the marginal value of production function. Then,

$$\frac{\partial^2 V\left(n, n_0, x\right)}{\partial n \partial n_0} = \frac{\partial^2 F\left(n, n_0, x\right)}{\partial n \partial n_0} > 0$$

where the parameters are properly chosen in the main text. Next, I assume that unknown and unproductive workers cannot contribute to the total production in the following derivation:

$$\begin{aligned}
F\left(n, n_0, x\right) &= x\left(\omega\left(\overline{y}n\right)^{\rho} + \left(1-\omega\right)\left(n_0\right)^{\rho}\right)^{\frac{\alpha}{\rho}} \\
&= x\left(\omega\left(\overline{y}n\right)^{\rho} + \left(1-\omega\right)\left(\left(Pr\left(y=\overline{y}\right)\overline{y} + Pr\left(y=\underline{y}\right)\underline{y}\right)n_0\right)^{\rho}\right)^{\frac{\alpha}{\rho}} \\
&= x\left(\omega\left(\overline{y}n\right)^{\rho} + \left(1-\omega\right)\left(\left(Pr\left(y=\overline{y}\right)\overline{y}\right)n_0\right)^{\rho}\right)^{\frac{\alpha}{\rho}} \\
&= x\overline{y}^{\alpha}\left(\omega\left(n\right)^{\rho} + \left(1-\omega\right)\left(Pr\left(y=\overline{y}\right)n_0\right)^{\rho}\right)^{\frac{\alpha}{\rho}}
\end{aligned}$$

where $\underline{y} = 0$. Therefore, the high level of production from known workers and "unknown and productive" workers show up the front in the production function, as functioning in a scale parameter. This is similarly defined in Elsby and Michaels (2013), which implies the mean production or fixed effect at the firm level.