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, Alessandra Peter, Junghum Park, Hyunjae Kang, Yurim Lee, Seungki Lee, Joonseok Oh, Sang-Ha Yoon, 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.

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 and Haltiwanger (2014). 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 11 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 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 growing literature on the nature of potential entrants and their post-entry behavior in studying firm growth and dynamics. Pioneered by Hopenhayn and Rogerson (1993), many studies try to find a source to explain a substantial difference between firms. Sterk *et al.* (2021) take ex-ante shocks into account in the Hopenhayn and Rogerson (1993) model, and Choi *et al.* (2019) study how the characteristics of business founders affect their growth. Unlike them, this paper has no ex-ante heterogeneity, but emphasizes post-entry endogenous decisions of firms changed with two ex-post shocks. Sedláček (2020) also studies post-entry behavior varied with aggregate conditions, but focus more on cohort effect between firms. Beyond the cohort effect, this paper tries to understand the impact of worker reallocation on firm growth, and its mechanism within and between firms.

A large body of empirical literature investigate 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) recently find a V-shaped pattern of excess hiring, as the same measure in this paper, associated with employment growth in the QWI. Bachmann *et al.* (2021) confirm a similar pattern in a German case. These results imply the fundamental relationship, beyond country-specific factors, between worker reallocation and firm growth. This paper supplements their findings and further, suggests a theoretical framework to explain its mechanism. Using the panel structure in the QWI, I track how young entrants grow over time by their worker reallocation over the first three years. Then, I find that their speed of reallocation is substantially unequal between fast and slow growing firms. This implies that a greater extent of firm-level heterogeneity exists within the V-shaped pattern.

The theoretical framework in this paper also contributes to another growing literature of the multi-worker firm model. The model of this sort explains how gross job flow differs by gross worker flow (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)). The main difference is that this paper considers excess hiring as a main driver in firm growth. In Bilal *et al.* (2019), fast-growing firms are always on the top of the marginal surplus distribution, then they do not have any incentive to reallocate workers beyond their job changes. Gross worker flow is tightly connected to the level of gross job flow. A firm in Elsby *et al.* (2020) creates excessive hiring, so called "vacancy chain," but it neither improves their match quality nor influences firm growth. Unlike the previous studies, this paper allows a firm to improve their match production by reallocating their employees every period, and this positive feedback comes into a benefit for the next period. Thus, fast-growing firms in this paper retain more productive workers to unknown workers. To this end, I put a similar learning process as in Jovanovic (1979), and the learning shock stimulates worker reallocation and size adjustment. A firm in this paper, therefore, optimizes not only their total size, but also the ratio of productive and unproductive workers over the life cycle.

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 private sector, but over 95% of jobs are collected in the database (Abowd *et al.*, 2006). As in Tanaka *et al.* (2020), this paper uses observations across 17 states in the US from 1990 to 2014.[4]

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.[5] Next, I compute establishment age from the calendar quarters when the establishment starts their business with non-negative number of employees, and drop out the establishment observations if they are initiated by older firms.[6]

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.[7] That is, the length of observed quarters varies

[4] 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)

[5] It is well-known fact that the distribution of employment size has a 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 hire and separation rate would be 2% and 1%, but if their average size was 2, then hire and separation rate would be 100% and 50%. To avoid this, I drop out *too small* employers as in Tanaka *et al.* (2020).

[6] 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.

[7] 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 $t+1$.

from 11 to 44 quarters.[8] Lastly, 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.[9] 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 the denominator follows 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.[10] The variable of this sort computes the probability of employees being replaced in their current employer rather than identify "who has been replaced by whom." The EAR in the main findings is averaged out over the first three years of firm age to track the initial average of worker churn.

The speed of churn, as represented by the EAR, depends on the distribution of firms for each categorical groups of firm age or so. For instance, 30% of the EAR can be fast enough for 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. Different categorical groups have a different level of fast or slow churners.

This approach has a couple of limitations. First, worker churn is time-varying and closely related to the job flow. A firm may replace their worker ahead of job creation or hire more workers for their delayed plan. The initial excess hire, by construction,

[8] Although the selection of this sort allows a short exit of firms, it is not completely free of impact from 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.

[9] 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)

[10] However, the EAR does not imply that the employee A in t has been replaced by a new employee B in t. 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.

reduces the bias of this sort. The rank of individual churners, faster or slower, is not substantially changed later once they positioned on the top of distribution. That is because worker churn is concentrated in young firms and it is declining as they age. In doing so, time-varying bias of worker churn does not have a significant impact on the analysis.

The worker churn cannot distinguish the case either 1) the employer loses productive workers (by demand shock or workers are poached) or 2) the employer actually separates unproductive workers. 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

How does firm size look like by worker flow? Figure 1 presents novel relationship, as guessed by Burgess *et al.* (2001), between worker churn and firm size. The centered gray line represents the average size of all firms. The marginal change in firm size decreases as firms age, consistent with a stylized fact in firm growth (Sterk *et al.*, 2021).

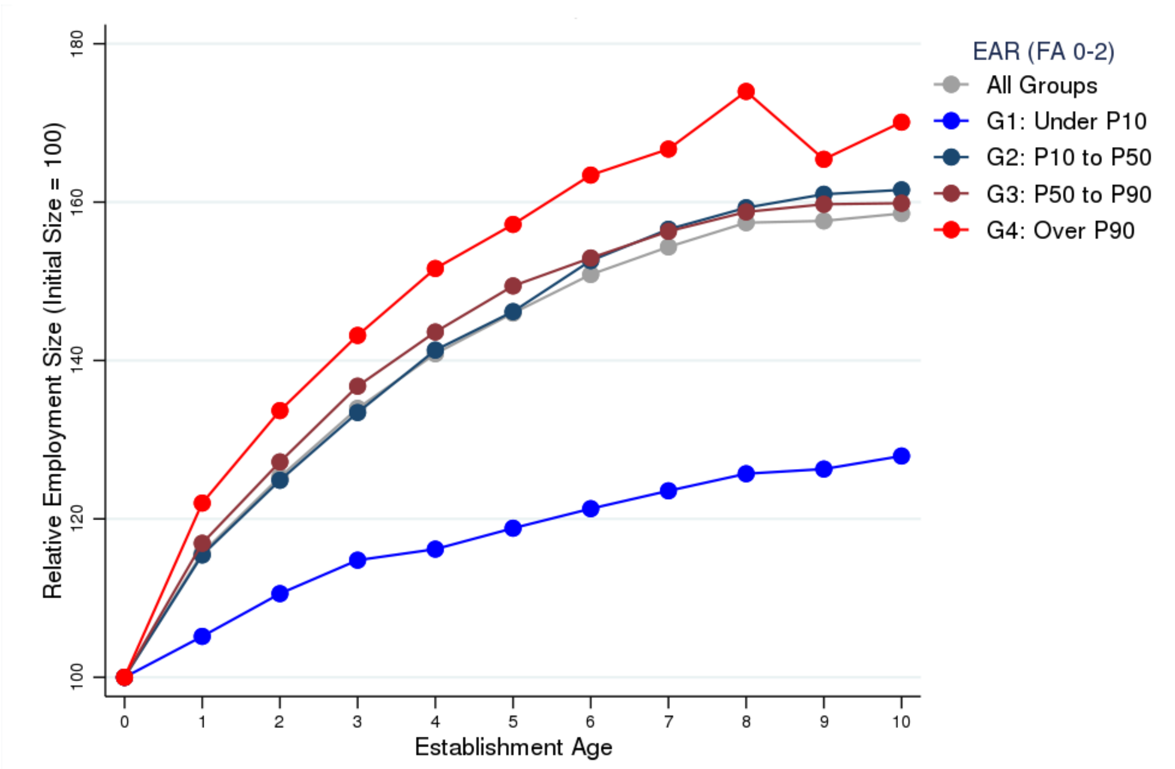
For the next step, I trace out their relative speed of churn when they are young, firm age between 0 to 2, and compute their average size later on. Firm size is denominated by their average size in the first year after entry. Conditional on the firm size, the blue line (G1) represents that a firm has worker churn in bottom tenth of overall distribution when young. These slow churners grow only 30% after 11 years of survival. The faster growing rate they have, the greater extent of worker churn is observed. The top growing firms, grow more than 60% to their size in their first year of age, are also top churners when they are young.

The gap between faster and slower churners widens and looks persistent. If the impact from worker churn faded away in the long run, the size disparity between firms is more likely to disappear as Acemoglu and Hawkins (2014) predict. But the size trajectory in figure 1 is far off from the prediction in the literature, rather shows a diverting pattern.[11] Weighted average by firm size does not change the overall trend. See the appendix for additional figures with weight and original size before denomination.[12]

[11] According to Acemoglu and Hawkins (2014), firm size differs by the ex-post heterogeneity such as adjustment cost, demand shock or so. Since all firms have a same target size in the end, they grow at unequal speed and end up the same place in the long run. This figure, one of the criticism, can be a snapshot in the transition period of this sort. Disparate nature before entry, like ex-ante shocks, might be one source of creating permanent differences outperformed post-entry behavior (Sterk *et al.*, 2021). However, rather than look for ex-ante heterogeneity, this paper tries to find a source to affect post-entry decisions. More details will be introduced in the theoretical section to account for this persistent difference.

[12] Cohort and industry fixed effect may affect this result. For cohort effect, Sedláček (2020) study how young entrants during the Great Recession have less potential to create jobs than other entrants in different

Figure 1: Trajectory of Firm Size by Initial Replacement Hires



Note: Each dot represents pseudo, unweighted average of firm size in each group of replacement hires in all industries. All numbers are denominated by firm's entry size.

What if big or small firms have substantially different worker churn after they enter? To check out the rank of worker churn by entry size, I compute the threshold of worker churn for categorical groups of firm size when firm age is zero. Figure 2 displays the average worker churn, for top tenth, median, and bottom tenth, over the size bins.[13] The

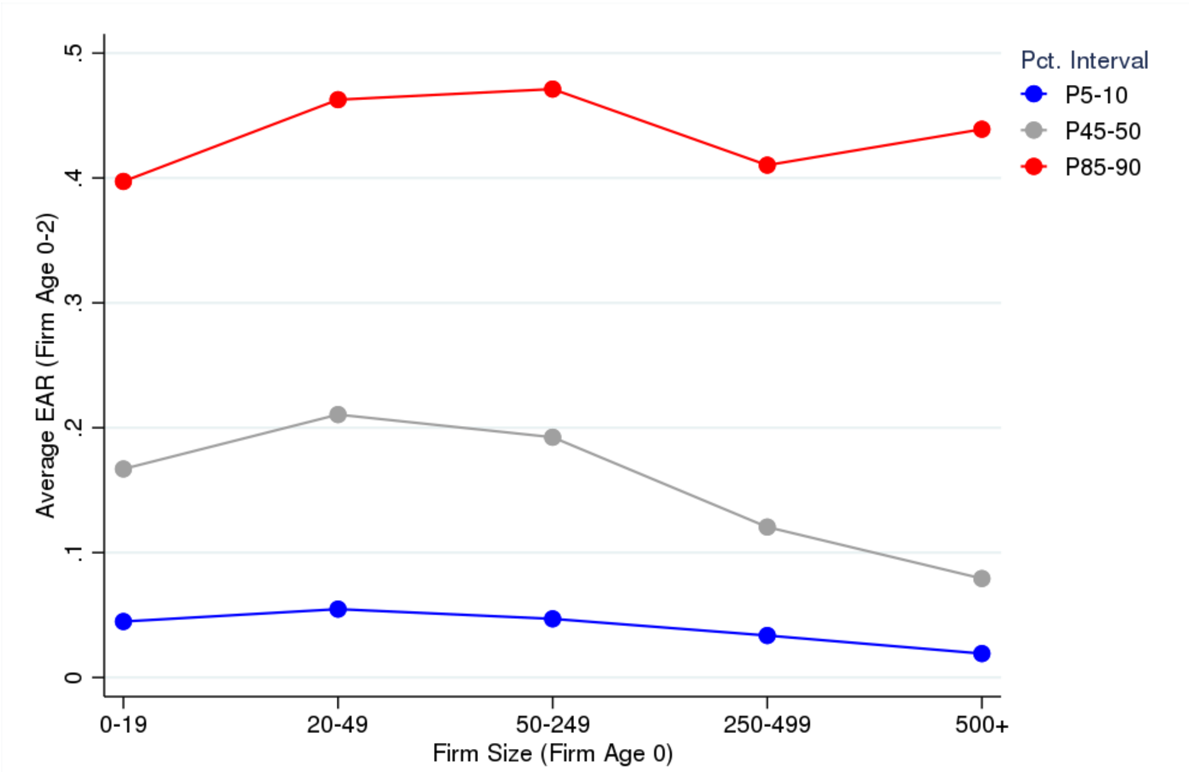
aggregate conditions. They particularly highlight that this aggregate fluctuation mainly drive the persistent differences across cohorts rather than ex-post decisions. Although gross firm entry and worker churn are both steeply declining during the Great Recession (QWI and BDS), there are still larger fluctuation of worker flow at the firm level. One paper to support this notion is Barrero *et al.* (2020). They find that many firms predict worker and revenue reallocation in their business survey. This proves that economic downturns (or aggregate conditions) do not necessarily harm worker reallocation. Moreover, I pool out all firms in one sample, so particular effect from the cohort in economic downturns is more likely to be canceled off by the other cohorts in normal times.

This paper only reports the economy-wide firm growth while there are still differences across industry. As in Kuhn and Yu (2021), it is well-known fact that wholesale or retail industry have a lower cost of worker turnover, so that their replacement flow is high. In the QWI, each industry has a different gross worker flow, so it is natural to think about the heterogeneity at the industrial level, and ex-ante differences of potential entrants. For instance, workers in wholesale industry are readily replaced by equivalent skill of workers while a manager in financial industry is not. Potential entrepreneurs may consider this cost when they enter, and subsequently this affects their post-entry behavior. I leave this industrial heterogeneity for future work.

[13]For confidentiality issue, I compute the pseudo average of the EAR between two percentiles. For

threshold of faster churners, represented by red dots, ranges from 40% to 48%. The cut point of slowest churners in blue dots also is not significantly different between small and big firms. The median of worker churn declines when firm size is greater than 250, so the distribution is slightly right-skewed. Overall, I conclude that the firm size in the first year does not substantially distort the distribution of worker churn.

Figure 2: Initial Distribution of Replacement Hires by Firm Size



Note: Each dot represents pseudo, weighted average of the EAR over firm age 0 to 2 in all industries. The interval of percentile is used to avoid the exact number of the average EAR in firm size groups. All numbers are weighted by firm size at age 0. The category of firm size follows the definition in Business Dynamics Statistics.

Back to the main question, what makes the firm size differ by worker flow and persistently divert between firms? The job creation process as in Hopenhayn and Rogerson (1993) or standard firm dynamics model does not fully account for the persistent difference because all differences from ex-post heterogeneity disappear in the long term. Also, their model does not allow excess hire or separation flow. In multi-worker firm model, as in Elsby *et al.* (2020) or Bilal *et al.* (2019), excess reallocation is not optimal due to the expensive adjustment cost. However, fast-growing firms, supposed to be

instance, bottom tenth herein is the average threshold between 5 to 10 percentile.

productive, have higher worker churn than slow-growing firms in the empirical findings. To explain this new pattern, this paper introduces a theoretical framework based on standard multi-worker firm model, combined with adjustment costs and employer learning in the next section.

4 Theoretical Model

I build a multi-worker firm model, with a learning process, in a frictional labor market. A firm in this model produces one single output with two different types of labor forces: known and unknown workers. While known workers are fully productive, the productivity of unknown workers is not 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." Then, the employer separates the unproductive workers and at the same time, determines the number of posting vacancies and separating incumbent workers for the gain and the loss from learning. The intensity of learning encourages a firm to reallocate their workers and drives a change in firm size. That means the composition of workers and total employment size are closely related in this economy.

In quantitative exercises, I will show that the mechanism of this sort well matches the key moment in the data set and accounts for the size difference between churners in figure 1. The persistence of employers' learning ability, creates the size of the gap between fast and slow churners. I simulate a similar transitional path between firms as in figure 1 with a different persistence of learning shock. Started by overall model environment, I will describe a firm's value problem with wage determination, and aggregate variables.

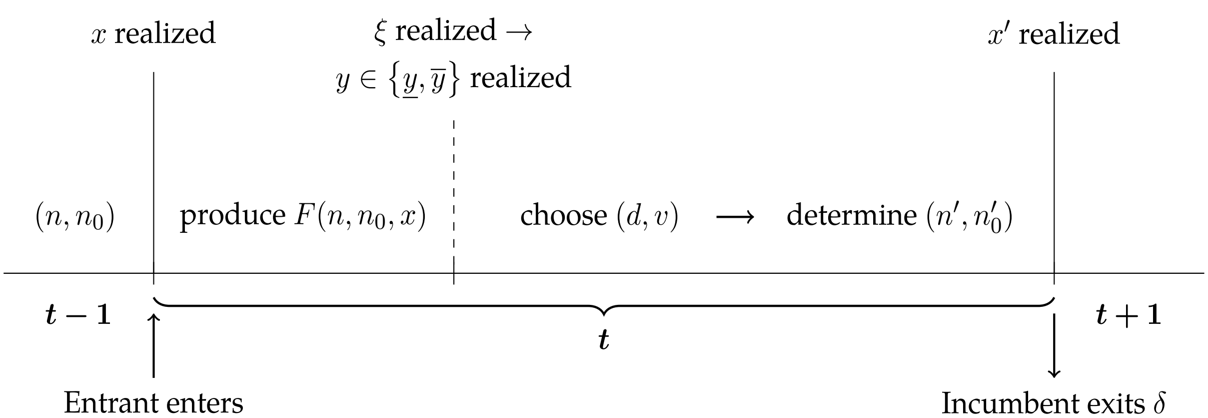
4.1 Model Overview

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 the fraction of them because their learning intensity is disproportional, implied by realized value of ξ. Employers thus have unequal speed of learning. In the meantime, unproductive workers are immediately separated and productive workers are promoted to a new group of known workers.[14]

Two policies are made after learning: vacancy positing (v) and endogenous separation (d). The firm posts vacancies to fill out the gap for the loss of unproductive workers from learning. Meanwhile, they consider separating incumbent workers because it is expensive to retain all labor forces. Combined with adjustment costs, the employer decides to hire and fire at the same moment. Not only their total size, but also the worker composition affects their future growth, so this is a new mechanism in this model. The incumbent firms exit at a constant rate (δ) before the next period comes in. Figure 3 outlines the infinite sequence of firm's decision and learning.

Figure 3: Model Timeline



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). A total size of labor force is $\mathcal{L} = E + \mathcal{U}$

[14] Silva and Toledo (2009) have a similar setup for labor forces, entrant and incumbent workers. In their model, a firm knows their entrant's productivity with a constant rate after employed. This paper takes a similar strategy for worker's productivity, but the learning intensity gives rise to the disparity between firms.

where E is employed workers and $\mathcal{U}$ is unemployed workers. Employed workers cannot search on the job until they are separated to the current employer. The aggregate vacancy posted is $\mathcal{V}$. The new match Z follows the matching technology with a 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 a new part in this multi-worker firm 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 employee's outside option, unemployment benefit, which will be explained in the next section. When a firm tries to change their size either by posting vacancies or firing their employees, they pay additional adjustment costs.[15] These two policies determine the number of workers employed in the next period, and pin down the flow profit.

Next, 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. Unknown workers who become productive are promoted to a part of known workers. The rest of unknown workers would be able to stay unless they are fired. Then, a firm has both size losses and gains from the learning of this sort.

In the amount of vacancies posted, a firm considers the size loss in unknown workers, but also the future gain of adjusting their total size. These new hires are added to unknown workers in the next period. Both known and unknown workers quit their employer at a constant rate s, so it is "worker-initiated" quit, while the endogenous separation is "employer-initiated" quit (Silva and Toledo, 2009). Therefore, the full de-

[15] If a firm has no frictions such as learning, adjustment costs, or ex-post shocks, they could jump to their optimal size by posting large amount of vacancies. Ex-post heterogeneity of this sort prevents a firm from the immediate changes.

scription of firm's problem is

$$V(n,n_0,x)=\max_{v,d\in\{d_1,d_2\}}\Big\{\underbrace{F(n,n_0,x)}_{\text{Production}}$$
$$-\underbrace{w(x)(n+n_0)}_{\text{Labor Cost}}-\underbrace{C(v)}_{\text{Vacancy Posting Cost}}-\underbrace{C(n,n_0,d)}_{\text{Separation Cost}}$$
$$+\beta(1-\delta)E_{x',\xi'|x,\xi}V\left(n',n_0',x',\xi'\right)\Big\} \tag{3}$$

subject to

$$n'=(1-d_1)\,n+\xi n_0\cdot Pr\big(y=\overline{y}\big) \tag{4}$$
$$n_0'=(1-d_2)(1-\xi)\,n_0+vq(\theta) \tag{5}$$

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

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 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 $d=s$.

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 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).[16] 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

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

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

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 the 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:

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

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.

4.6 Aggregation

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

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

As in Elsby and Michaels (2013), the job creation and Beveridge curve determine the number of unemployed workers in equilibrium. In 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 cannot find a job in this period. Those two workers enter the next period as a new mass of total unemployed.

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

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.

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

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

From the definition of 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.[17]

5 Calibration

I calibrate the model to the US economy. The model frequency is monthly to replicate a 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 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 a 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 $q(\theta) \fallingdotseq 0.5556$.

[17] 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.

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.[18] I add the decreasing returns to scale to pin down their total size in equilibrium.

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.[19]

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

where $\rho = \frac{\sigma-1}{\sigma}$, $0 < \sigma < 1$ and $0 < \omega < 1$.[20]

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 total number of establishment, is close to 10%. Then, the monthly exit rate δ is 0.83%, divided by 12 months.

---

[18] 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

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

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

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.

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

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

A constant fraction of workers is separated from their current employer in every period. This exogenous outflow is set to match the quit share of 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 Mortensen-Pissarides search model, the reservation wage level equals to 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 $b = 1 - \beta(1-\delta) = 1 - 0.996 \times (1 - 0.0083) = 0.012267$.

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

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

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:

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

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

In Silva and Toledo (2009), the firing cost in the US is estimated to 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 of 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 $\varepsilon_x' \sim N(0, \sigma_x^2)$. 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 persistence of this shock, $\rho_x = 0.963$, but choose σ_x^2 to prevent a negative flow profits across all the states in 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.[21] I assume the learning

[21] 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 center accounts for 15% to 22% out of total separation in one quarter (Bureau of

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.[22]

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

where $\varepsilon'_\xi \sim N\left(0, \sigma_\xi^2\right)$. As the grid comes in for the simulation, a different rate (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 degree of 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 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 showed before, determines the slope of optimal path in frictionless setup. The weight parameter, 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 job in Hyatt and Spletzer (2017). They report that the average rate of single quarter job is 8% and the multi quarter job is 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 current and previous period. 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 of correlation between log employment is 0.852, which is highly persistent.[23]

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,

Labor Statistics, 2012). This data set provides the number of separations by categorical reasons, but none of them directly links to the learning intensity. In fact, learning may exist in all those categories. Another source is the Displacement Worker Survey, reported every three year for supplements of the Current Population Survey. This also offers specific reasons for 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%) accounts for 91% of total displacement in all industry (Kandilov and Kandilov, 2010). All those could have displaced workers initiated by employer learning.

[22] The joint pair is $(\rho_\xi, \sigma_\xi) \in \{(0, 0.15), (0.5, 0.13), (0.75, 0.1), (0.9, 0.06)\}$ and this implies $\xi \in \{0.2, 0.35, 0.5, 0.65, 0.8\}$.

[23] 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.

within three months.[24] 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 job 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 to the job creation rate as in Borovickova (2016). In the BDS, the average of job creation rate from 1990 to 2014 is 15.3%. 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.

[24] I follow the notation from Steven Stern's lecture note for survival analysis.

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		Quarterly share of mass layoff
σ_ξ	AR(1) learning std.		$\xi \in \{0.2, 0.35, 0.5, 0.65, 0.8\}$
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 job in total = $\frac{1}{3}$
c_1	Scale parameter	0.400	Quarterly job creation rate = 3.825%

6 Model Outcomes

6.1 Targeted and Non-Targeted Moment

Table 2 summarizes the model performance. In baseline case ($\rho_\xi = 0$), the model fits the data well except for the job creation rate.[25] 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 the model is almost equivalent to the target in the data.

The bottom in the same table display the result of non-targeted moment 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 groups.[26] The annualized growth

[25] 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.

[26] Firm size of churners in G1 ends up 30% greater than their entry size in their age 10, so the annualized

rate between fastest and slowest churners implies the 23.5% of size gap. The baseline model has a smallest difference in the average growth rate between firms, so 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.

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

6.2 Simulation Results

I simulate 25,000 firms for 11 years of firm age, 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.[27]

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 equals to zero) is 10.7. Motivated by this empirical moment, all firms at age 0 employ a half of the average size of incumbent firms.[28] Suppose total size is the sum of known and unknown workers ($N = n + n_0$).

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 $1 - \frac{130}{170} \fallingdotseq 1 - 0.7647 = 0.2353$.

[27]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.

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

Then,

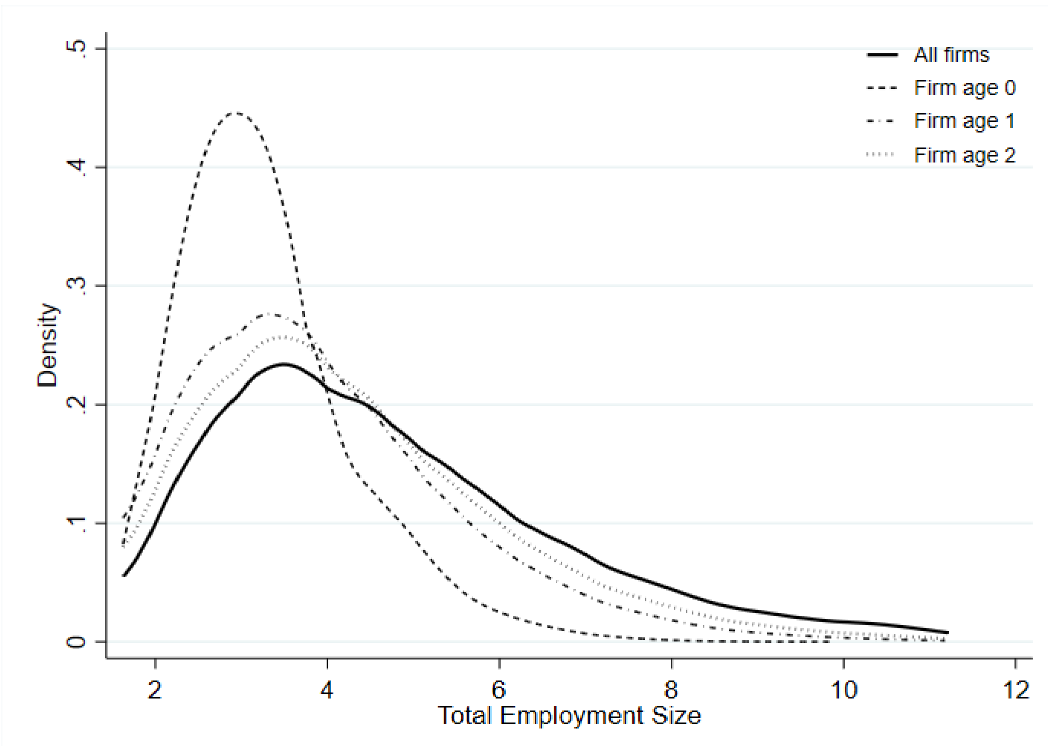
$$N_{\text{FirmAge}=0,t=1} = \frac{1}{2} \times \overline{N}$$

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 worker 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.[29] The figure 4 shows that the size distribution of young entrants (dotted lines) is rightly skewed than incumbent firms (solid line) and their average size increases as firms age.[30]

Figure 4: Size Distribution of Entrants and Incumbents



Note: This graph displays kernel density estimates using Epanechnikov algorithm for the case of $\rho_\xi = 0.75$. I select 0.5 for bandwidth.

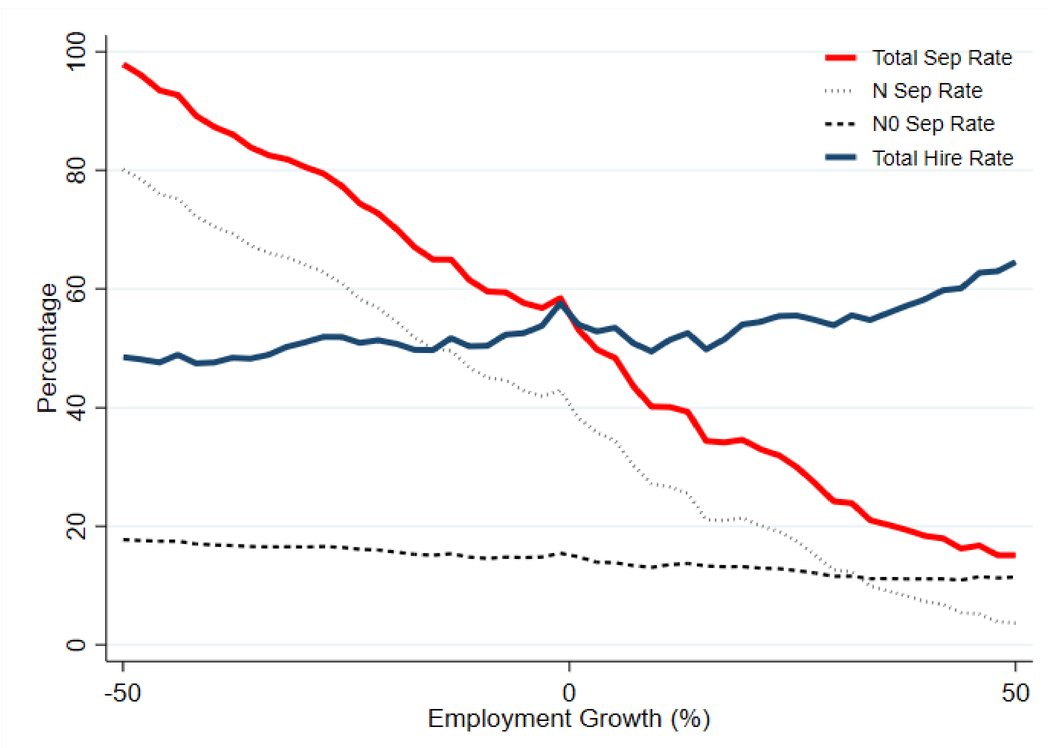
Figure 5 introduces a cross-sectional relationship between worker flow and employ-

[29]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.

[30]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.

ment growth. The Expanding firms have a higher hiring rate than separation while contracting firms show the opposite. Davis and Haltiwanger (2014) report the similar pattern between hire and separation in the data set.[31] Gross change in worker flow directly implies the net change in job flow. This graph has more lines for the number of known and unknown workers separated from contracting or expanding firms. For instance, contracting firm separate known workers more than unknowns, which implies 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 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.

Figure 5: Worker Flow over Employment Growth



Note: This graph displays the cross-sectional relationship between worker flow and employment growth in firm simulation for the case of $\rho_\xi = 0.75$. The range of the horizontal axis is restricted from -50% to 50%.

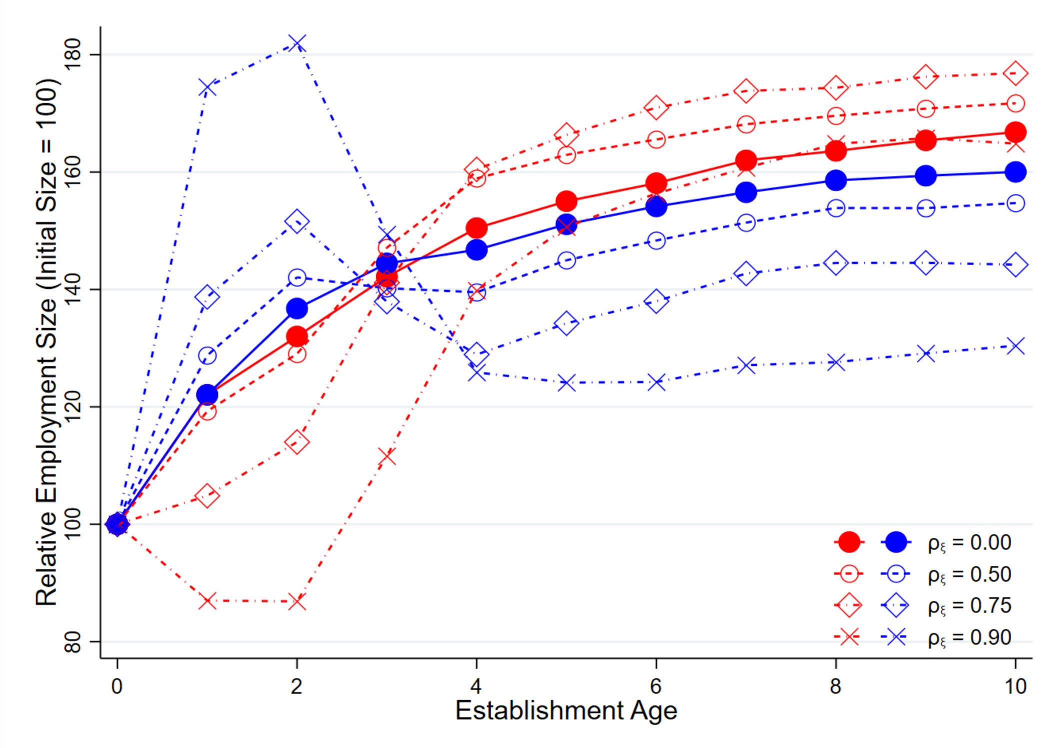
[31] This paper does not have a kinked point as their plot, so called "hockey stick." In the model economy, a firm separates many workers by learning and endogenous allocation. Future hires, given the vacancy filling rate, consider all these separations as in a mirror relationship.

6.3 Source of Persistent Difference

The sizable gap in figure 1 does not disappear after 11 years, so it is crucial to understand its source of persistent difference. From the baseline estimates, I simulate the model economy with a different extent of persistence in learning. The figure 6 displays the transitional path of fastest and slowest churners for each persistence in the same years of activation. If a firm does not know their next learning shock, which is $\rho_\xi = 0$ and only i.i.d. part remains, the gap between the two lines has almost disappeared. This is consistent with the outcome of non-targeting moment in table 2. In this case, their size gap is only 6% while the data shows 23.5%.

But the greater extent of persistence creates a sizable gap as in the empirical finding. From $\rho_\xi = 0.5$ to $\rho_\xi = 0.9$, the difference between the red and blue line after 5 years becomes larger. According to the table 2, the increasing persistence explains 30-60% of the size gap in firm age 8 to 10. Another interesting point is that there is a size overshooting over the first 4 years. The slowest churners momentarily outgrow than fastest churners because a firm can avoid adjustment costs if they only lay off a few workers in the first place. But soon their size is declining due to the negative productivity effect from unproductive workers. In the meantime, fastest churners downsize or are slow growing for a while because of the unavoidable cost of adjustment for their policy choices. Once they overcame this period, they grow on the top of the size ladder. These two examples how employer learning has an influential impact on firm growth in this economy.

Figure 6: Relative Size and EAR



Note: The red line represents top tenth (fastest churners) and the blue line is bottom tenth (slowest churners) in the distribution of worker churn over firm age, which is equivalently created as in figure 1.

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 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. The 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 simulation, I show that expanding firms have a better composition of their workers than shrinking firms. Also, the greater extent of 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 the 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 grow 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 worker's wage growth or employment status in job spells.[32] The decomposition of this sort enables to see how a firm replaces their workers with micro-level heterogeneity as in Kuhn and Yu (2021). If a firm shows a higher worker turnover due to the job-to-job transition, the firm is less likely to grow later on. A set of reasons for separation also helps understand whether employers separate their workers by learning or other reasons.

Ex-ante differences of potential entrants, pioneered by Sterk *et al.* (2021), is another interesting direction. Potential entrants in wholesale industry tend to hire lower skill of workers than in finance industry when they enter. Worker churn, with the ex-ante disparity, may have an uneven effect in 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.

In this paper, all workers are not allowed to search on the job and paid evenly regardless of their productivity level. In the future, I may consider the wage bargaining protocol as in Elsby and Gottfries (2019), Elsby *et al.* (2020) and Brügemann *et al.* (2019). They have developed a bargaining wage in multi-worker firm model allowing for on-the-job search. In doing so, productive workers have a better wage scheme than unproductive workers in terms of their outside options and marginal contribution to the total production.

[32] 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 worker's earnings growth. This only looks at the change in 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.

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. 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. (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)

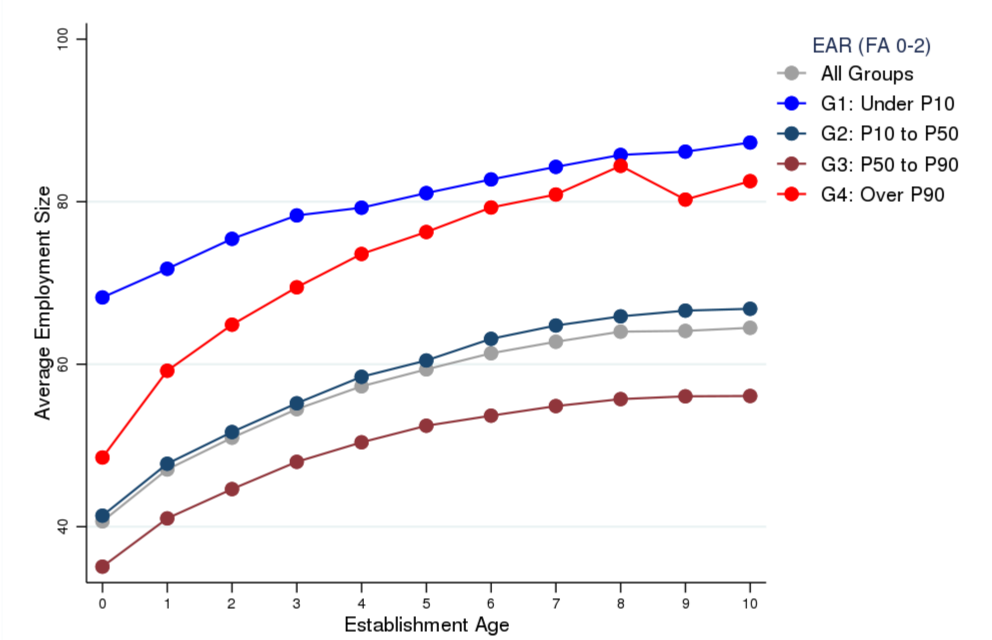


Figure A.2: Trajectory of Average Firm Size by Initial Replacement Hires (All Industries, Weighted)

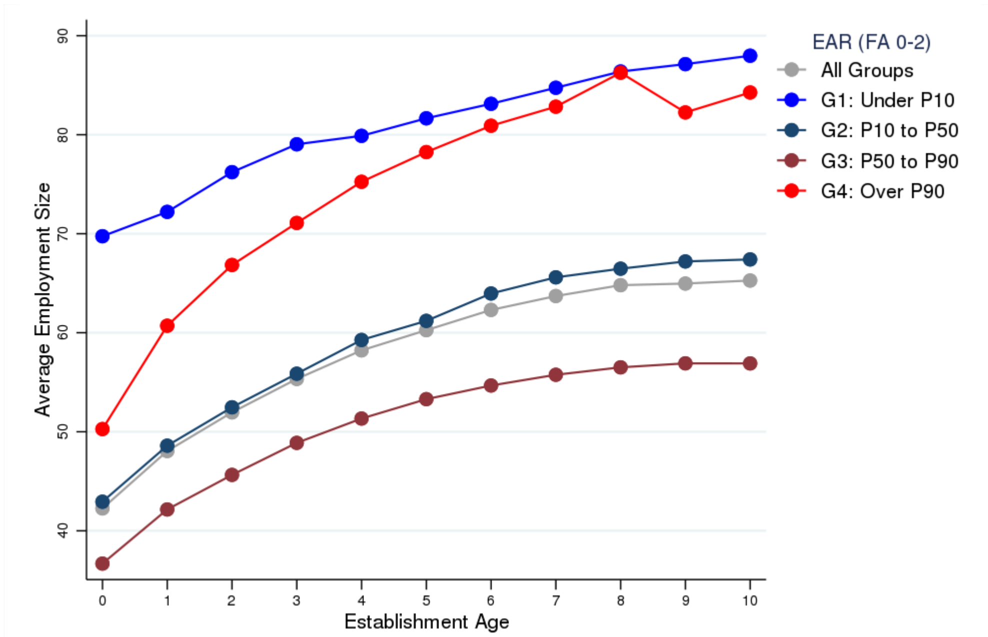
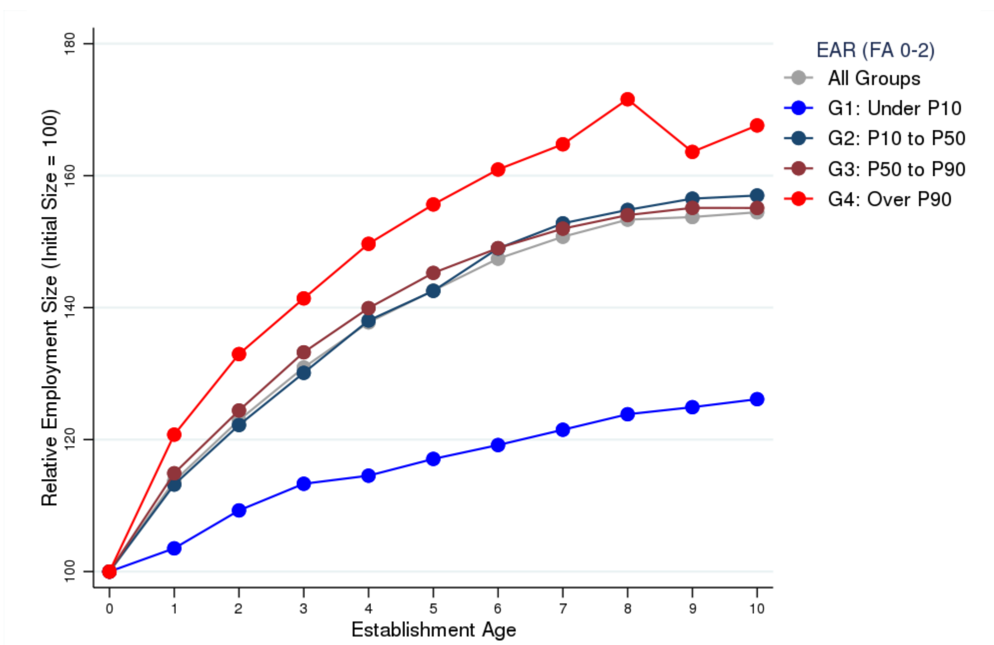


Figure A.3: Trajectory of Relative Firm Size by Initial Replacement Hires (All Industries, Weighted)



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:

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

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 functioned 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.