Population Aging, Firms, and Worker Composition

Santiago Caicedo, Cheng Chen, Aspen Gorry and Takahiro Hattori*

March 2026

Abstract

Population aging changes the *composition of workers* in the labor force. We develop a framework to understand how worker composition affects productivity in an economy with heterogeneous firms and apply it to Japan. Reallocating workers toward the optimal composition of experience and skill would increase output by 3.8%. Including worker aging and reallocation costs, we find strong complementarity between workers' skill and firm productivity, which generates differential hiring strategies across firms and heterogeneous effects of population aging on wages and market concentration. Using Japanese establishment data, we validate key model predictions: larger firms disproportionately hire young workers, and regions with slower working-age population growth exhibit lower concentration of employment among top establishments.

Keywords: Population Aging, Labor Force Composition, Comparative Advantage of Productive Firms

JEL-Codes: D21, D22, E23, E24, J11, J23

*Caicedo: Northeastern University, sa.caicedo@northeastern.edu; Chen: Clemson University, cchen9@clemson.edu; Gorry: Clemson University, cgorry@clemson.edu; Hattori: The University of Tokyo, hattori0819@gmail.com. We thank Sadhika Bagga, Seula Kim, Victoria Gregory, Elena Pastorino, Ben Pugsley, Michael Wong for insightful suggestions and the audiences at the 2023 Southern Economic Association Meeting, the University of Hong Kong, the 2025 ASSA annual meeting, the Barcelona Summer Forum, Bentley Applied Macro Workshop, and Northeastern University for their helpful comments. All errors are ours. We thank Isabella Becerra for her excellent research assistance.

1 Introduction

The composition of workers by observable demographic groups, such as experience and skill, is a crucial factor in determining aggregate economic outcomes. For example, increases in cohort size of the baby boom generation reduced wages and wage growth for these larger cohorts, highlighting that workers with different experience are imperfect substitutes in the labor market (Welch, 1979; Berger, 1985). Moreover, if heterogeneous firms employ different portfolios of workers in production, compositional changes in the workforce can have differential impacts across firms. Recently, the composition of workers has been evolving rapidly due to population aging, declines in fertility, and immigration (Toossi and Elka, 2017; Cardona, 2023; Frey, 2025; Qi, 2025; Geruso and Spears, 2026). Given large differences in demographic patterns across countries and over time, it is important to understand how changes in labor supply composition influence the allocation of workers across firms and aggregate productivity.

This paper seeks to characterize the optimal allocation of workers to firms, using Japan as a case study to understand the optimal worker composition and the impacts of aging. Japan is an interesting case because it is a country at the forefront of demographic aging.[1] Japan also has establishment data that contains rich worker and demographic information, including worker tenure. Using this data, we first document that different establishments do indeed hire different portfolios of workers, with larger (more productive) establishments having younger workers on average but higher average tenure. This implies that large establishments have steeper experience-tenure profiles than smaller ones, suggesting they hire a disproportionate number of labor-market entrants. Understanding the optimal portfolios of workers across establishments provides evidence on the optimal mix of workers in production and informs policies that influence the composition of demographic groups, such as immigration, education, fertility promotion and retirement programs.

Motivated by these facts, we develop a framework to characterize the efficient composition of workers across heterogeneous establishments. We start with a static framework that highlights how labor composition across groups influences labor allocation and aggregate output. The economy consists of heterogeneous establishments that vary in productivity and employ workers from discrete groups. The key assumption is that labor markets clear at the group level, so different firms all pay the same wage to hire workers from each group. This framework allows us to characterize the optimal allocation of workers across firms given a distribution of workers by group. Moreover, given the production technology, the model also characterizes the optimal composition of workers for production.

[1] The median age of the Japanese population was 48.4 years in 2020 and has been increasing over time (IMF, 2020). Moreover, the World Bank reports that the growth rate of the working-age population has declined from 0.24% in 2001 to -0.45% in 2021. For recent work studying the impact of rapid aging in Japan, see Braun et al. (2009), Braun and Joines (2015), and Kitao (2015).

The optimal allocations described by the static model are consistent with the data if there are no costs of reallocating workers across establishments. In the presence of reallocation frictions, it is necessary to model both the cost of worker reallocation and the transitions workers make across demographic groups over their life cycle. We extend the model to include these dynamic features to understand worker transitions across groups and the importance of reallocation costs. The dynamic model features workers whose experience evolves deterministically and skills that evolve stochastically over their life cycle. Heterogeneous firms that differ in their productivity hire a portfolio of workers in each experience-skill bin and must plan for how that portfolio evolves over time. Finally, we include a firing cost that makes it costly to reallocate workers and distinguishes between internal and external labor markets, thereby providing the model with a meaningful definition of tenure.

With the static model as a benchmark, we use Japanese establishment-employee matched data to parameterize the model.[2] To provide evidence on group-specific labor productivity, parameters of the model are chosen to match key moments obtained from the microdata, including wages and the percentage of workers in each experience/tenure bin. Moreover, the parameter for the complementarity between firm productivity and skill is informed by matching the differential experience-tenure profiles across firms. The calibrated parameters imply strong complementarity (i.e., log-supermodularity) between firm size (i.e., productivity) and workers' skill, leading to differential hiring/firing strategies across establishments. High-productivity establishments hire disproportionately more young workers and fewer middle-aged and older low-skilled workers. These predictions are validated using information on hiring at the establishment-year level, showing that large establishments hire disproportionately more job-market entrants.

Using these group-level productivities, we compute the optimal composition of workers across groups within the model. Our main result is that the optimal composition has a larger share of experienced workers (those with 21-35 years of experience or 41-55 years old) than Japan's current population. Quantitatively, Japanese GDP per worker would be 3.8% higher with the optimal experience and skill composition. These gains are larger if the production function exhibits less diminishing marginal productivity for each worker group or if workers are more substitutable.

We re-estimate the results using a dynamic model with positive firing costs. In this environment, productive establishments have weaker incentives to hire young workers and stronger incentives to retain older, low-skill workers. As a result, young workers earn lower wages, and high-skill workers also experience wage declines. These changes lead to greater wage inequality across experience groups but lower wage inequality across skill levels.[3] Moreover, the presence of firing costs amplifies the estimated productivity differentials between high- and low-skilled workers among experienced

[2] Our approach takes the measured technology as given when considering changes in demographics. Modeling endogenous technological change driven by changes in worker composition is beyond the scope of our framework.

[3] Young workers earn less than older workers on average, both in the data and in the model.

workers. Relative to the benchmark counterfactual, GDP per worker increases by more when the economy is endowed with the optimal worker composition (4.8% versus 3.8%). Importantly, nearly all of this gain is driven by the recomposition of worker skills and experience, as eliminating firing costs alone raises GDP per worker by only 0.1%.[4]

Finally, we apply the model to understand how population aging in Japan influences aggregate output, wage inequality, and firm concentration. We simulate the evolution of Japan's demographic structure over the next 40 years, assuming the current fertility rate remains at 1.3. To isolate composition effects, we hold the total workforce constant throughout the simulation. Population aging has modest effects on aggregate output, with a maximum decline of around 1% over the 40-year period.[5] This result reflects two factors: workers with 21-40 years of experience are revealed to be the most productive, and the share of very old workers remains moderate even after 40 years.[6]

Population aging also affects wage inequality and market concentration. First, the wage premium of high-skill workers increases substantially over time, as workers mainly learn when they are young (with less than 10 years of experience), and population aging reduces the relative supply of high-skill workers. For firms, population aging hurts high-productivity establishments, as their comparative advantage lies in utilizing workers' skills rather than experience. A shift in the distribution of workers toward older workers raises the skill premium, hurting high-productivity establishments. As a result, their employment/output shares in the economy shrink over time. Beyond these model predictions, we use evidence from Japanese prefectures to study how population aging relates to firm concentration. We confirm that the employment and sales shares of the top establishments increase more when the growth rate of the working-age population and/or the average age of the population in that prefecture is higher.

Related Literature Understanding how firms choose a portfolio of workers contributes to several important areas of research. First, we relate most closely to a literature studying the impact of population aging on economic outcomes. Most closely related to our paper are papers that consider the effect of population aging on firm dynamics, including Peters and Walsh (2021), Hopenhayn et al. (2022) and Karahan et al. (2024).[7] These papers study how the growth rate of the size

[4]The social planner does not face firing costs when recomposing the distribution of worker characteristics to maximize output, since firing costs are not technological constraints. As a result, in the economy with firing costs, the recomposition gain—defined as the change from the decentralized equilibrium to the optimal distribution of worker characteristics—includes both the gain from eliminating firing costs (holding the distribution of worker characteristics fixed) and the gain from recomposing worker skills and experience in the absence of firing costs.

[5]The maximum reduction in aggregate output for the 40 years is around 1%.

[6]By the end of the 40 years, the fraction of workers with more than 31 years of experience is around 35%.

[7]Other papers on implications of population aging that do not focus on firms (or matching between workers and firms) include Maestas et al. (2016), Hayashi (2023), Bianchi and Paradisi (2024), and Weil (2026). This literature largely focuses on the mechanism by which a higher share of retirees (the dependency ratio) lowers per capita output and consumption, rather than on how changes in the age structure of the working-age population influence

of the labor force matters for firm entry/exit (i.e., dynamism), innovation, and labor share. Our paper complements this work by abstracting from overall changes in the share of the working-age population to focus on how changes in the composition of the labor force are important for firm productivity and inequality. Specifically, we show how an increase in the relative supplies of experienced and inexperienced workers affects the output of heterogeneous firms and market concentration.

Second, we relate to the literature on firm dynamics with worker heterogeneity. Recent papers in this literature, including Engbom et al. (2019), Gregory (2020), Bagga (2022), and Kim (2023), focus on workers who are heterogeneous in a single dimension: skill. We enrich this framework by emphasizing the importance of worker experience in the firm's optimal employment portfolio. Experience is interesting as it ties firm portfolios to deterministic patterns of demographic change driven by population aging. Our paper complements this literature by showing that firms have *heterogeneous* exposure to worker skills and experience because more productive firms hire relatively more skilled workers. This heterogeneous exposure has important implications for how population aging and firing costs influence aggregate productivity, product market concentration, and wage inequality between experienced and inexperienced workers.

Finally, our paper contributes to the literature on the impact of firing costs on firm outcomes. Bentolila and Bertola (1990) and Hopenhayn and Rogerson (1993) are the two key early contributions. Bentolila and Bertola (1990) find that the firing cost has larger effects on the propensity to fire than to hire, increasing the average employment of a single firm in a partial equilibrium setting. In a general equilibrium setting, Hopenhayn and Rogerson (1993) shows that firing taxes distort firms' hiring and firing decisions, decreasing output and productivity for the aggregate economy.[8] We introduce additional heterogeneity in firms' employment decisions and show that firing costs have important distributional effects on workers and firms.

The rest of the paper is structured as follows. Section 2 describes the Japanese establishment-worker data and key empirical findings. Section 3 develops the theoretical framework. Section 4 presents the calibration and model fit. Section 5 quantifies efficiency gains from optimal labor recomposition. Section 6 applies the model to study population aging and provides empirical validation. Finally, Section 7 concludes.

productivity.

[8]Subsequent research extends Hopenhayn and Rogerson (1993) by introducing capital into the production function (Veracierto, 2001) or search and matching frictions into the labor market (Alvarez and Veracierto, 2001) or innovation activities into the firm's problem (Poschke, 2009).

2 Worker Composition in Japanese Establishments

In this section, we describe the Japanese establishment dataset and present our main empirical finding: larger establishments have younger workers with longer tenure. Japan is an ideal setting to consider changes in the workforce composition, as it currently has one of the oldest populations with a declining share of the working-age population. Moreover, Japanese data are useful because the establishment-level employment survey provides information on worker tenure, so we do not need a long panel to construct an accurate tenure measure.

2.1 Japanese Establishment Data

We use the Basic Survey of Wage Structure (BSWS), a survey of the wage structure at the establishment level. The BSWS is conducted annually by the Japanese Ministry of Health, Labor, and Welfare. The universe of the survey is private establishments with five or more regular employees and public establishments with 10 or more regular employees in Japan, except those classified in agriculture, forestry, fishery, and the legislative, administrative, and judicial branches of local and national governments, with a new random sample of workers each year. The survey has been conducted continuously since 1948; however, we have access to data from 1998 to 2011. Approximately 55,000 establishments provide valid responses annually, and those surveyed include approximately 1.28 million employees. Although board members are not included, all types of workers (regular and temporary, full- and part-time) are surveyed after 2005. The survey provides detailed information on hours and wages, including overtime hours and annual bonus pay, as well as workers' experience, education, tenure, and work type (manufacturing/management/administrative/R&D). There is no top or bottom-coding of wage records.[9]

We link establishments over time using the identifiers provided by the Ministry of Internal Affairs and Communications. Specifically, the BSWS sample is derived from the Economic Census, which contains all establishments in Japan, each with a unique identifier.[10] This allows us to link observations of establishments that appear in the BSWS in multiple years, generating a panel of establishments from 1998 to 2011. We also have information on establishment age from the Economic Census for two sub-periods of our sample: 1998-2001 and 2004-2011.

Although the BSWS provides information at the worker level, we aggregate worker-level information

[9]The survey was redesigned in 2005 for the purpose of obtaining better information on non-regular workers and workers with fixed-term contracts. Before 2005, a surveyed worker only reported information on whether he/she is a temporary (or permanent) worker, where being temporary meant the length of employment was less than a month. After 2005, a permanent worker who is surveyed reports whether he/she is a regular worker and whether he or she has a fixed or open-ended contract.

[10]The unique identifiers of establishments in the Economic Census are revised every three to five years. We use the concordance between the current and previous identifiers each time the identifiers are revised.

Table 1: Summary statistics: Worker characteristics at establishment level (2005–2011)

	Mean	SD	P10	P25	P50	P75	P90	N
Employment	90.14	293.2	8	13	26	63	164	377,740
Share non-regular workers	0.663	0.338	0.103	0.381	0.788	0.984	1.000	377,740
Age (years)	41.33	7.891	31.11	36.44	41.33	46.24	51.33	377,740
Experience (years)	21.76	8.015	11.71	16.28	21.43	26.67	32.20	369,107
Tenure (years)	9.718	6.050	2.600	4.857	8.824	13.57	18.14	377,740
Education (years)	13.20	1.402	11.67	12.05	13.00	14.22	15.26	369,107
Monthly wage	290.6	135.4	118.3	187.0	284.1	375.2	468.4	377,740
Bonus	55.82	61.93	0.00	6.96	35.45	86.70	143.3	377,740
Total compensation	346.4	187.7	123.1	200.7	322.8	457.0	602.3	377,740

Notes: Summary statistics are calculated as follows: for each establishment-year observation, we compute the average of each variable across all workers in that establishment; the table reports the distribution of these establishment-level averages across all establishment-years in the sample. Experience equals age minus years of education minus 6. Monthly wage, bonus, and total compensation are expressed in thousands of yen (approximately 10 USD at 2005–2011 exchange rates). Non-regular workers include part-time, temporary, and contract workers.

to the establishment-year level, as workers cannot be linked over time. Specifically, we calculate establishment-level averages of worker variables and present their summary statistics in Table 1. In our sample, establishments on average employ 100 workers and have a mean age of 41. Japanese workers have a median tenure of 8.8 years. Japanese workers have significantly longer tenure than those in the U.S., where the median tenure was 4.2 years in 2016, according to the Bureau of Labor Statistics. The average years of schooling are slightly more than 13. The table also reports the annual salary, defined as the sum of the wage and bonus.

2.2 Larger Establishments Have Younger Workers with Longer Tenure

We next examine how these workers are distributed across establishments of varying sizes. Panel (a) of Table 2 documents that larger establishments employ workers who are slightly younger on average but have longer tenure. Workers in larger establishments are about 0.4 year younger (or have about 0.7 year less experience) than those in smaller establishments.[11] In contrast, these workers have roughly two more years of tenure with their current employer. This is surprising, as younger workers typically have shorter tenure.

Panel (a) of Figure 1 plots experience-tenure profiles for the two establishment size groups, illustrating that this pattern generates differential experience–tenure profiles across establishments. The experience–tenure profile is much *steeper* for workers in larger establishments. While young workers

[11] Experience is measured as age minus years of schooling minus 6. We do not account for unemployment spells, as this information is unavailable at the worker level, but this is unlikely to be a major concern given Japan's low unemployment rate.

Table 2: Average Worker Characteristics by Two Size Bins (2005–2011)

Panel (a): All workers and all establishments						
Size Bin	Employment	Age	Experience	Tenure	Wage	Education (years)
1	23.77	41.14	21.99	10.88	34606.24	13.15
2	404.98	40.70	21.22	12.84	40410.37	13.47
Panel (b): Workers younger than 61 and establishments established before 1985						
Size Bin	Employment	Age	Experience	Tenure	Wage	Education
1	24.40	40.55	21.34	12.10	35687.55	13.21
2	463.25	40.33	20.88	14.07	41599.32	13.45

Note: Time period: 2005–2011. Unit of average wage: one hundred yen (or one USD). We divide establishments into two size bins. Size bin 1: employment ≤ 50; size bin 2: employment > 50.

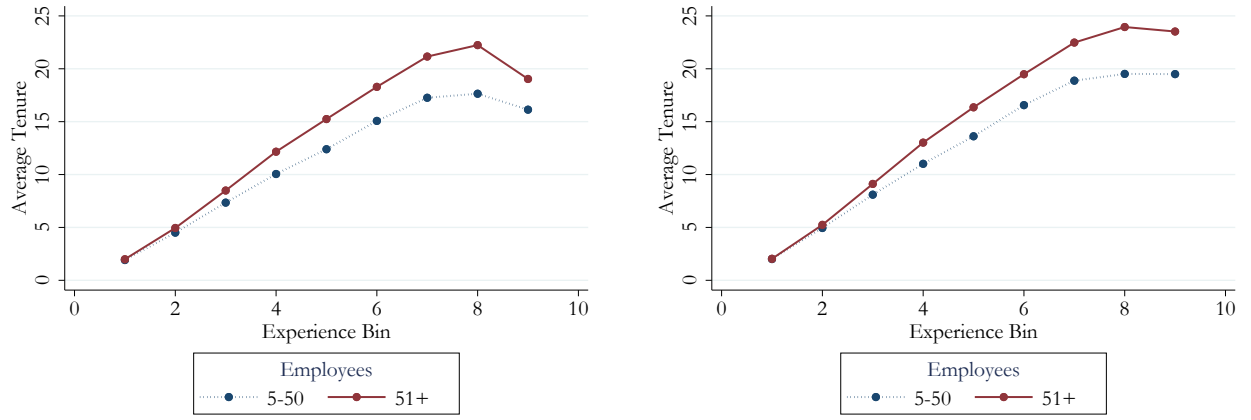
have similar tenure across establishments, middle-aged and older workers in larger establishments have substantially longer tenure. For example, workers with 36–40 years of experience in the larger establishments have about 22 years of tenure, compared with just 17 in the smaller establishments.

These findings are related to the finding that workers in large firms have lower separation rates than those in small firms in the U.S. (Papageorgiou, 2018). However, if small and large firms hire workers from a common age distribution, differences in separation rates alone are not enough to generate these facts, as the longer tenure at large firms would then imply that they have older workers on average. Hence, our findings suggest that large firms tend to hire more young workers as they enter the labor market, whereas smaller firms tend to hire more experienced workers. Bagga (2022) provides additional support for this mechanism in the U.S. context by showing that larger firms have a stronger internal job ladder. This is consistent with young, high-skilled workers being more likely to remain in large firms.

The pattern documented in Table 2 and Figure 1 are robust to different sample restrictions and measurement choices. First, the results are robust to restricting the sample over time and worker's type. Tables B.1 and B.2 and Figure A.1 in the appendix substantiate that the tenure–experience patterns become more pronounced when restricting the sample to the 1998–2004 period or to permanent workers. Second, Tables B.3 and B.4 and Figure A.2 in the appendix show that the tenure–experience patterns hold when we divide establishments to five equal-sized bins.[12] Third, the results are robust to measuring age instead of experience. Figure A.3 in the appendix demonstrates that the differential experience–tenure profiles across establishments hold when using age instead of imputed experience.

[12]For the period of 2005-2011, the main differences exists among three groups of establishments: size group 1, size groups 2-4 and size group 5.

Figure 1: Experience-Tenure Profiles (2005–2011)



(a) Full Sample

(b) Exclude Old Workers and Young Establishments

Notes: Experience is defined as age minus (length of education + 6). Experience bins are defined as: bin 1 (1–5 years), bin 2 (6–10 years), ..., bin 8 (36–40 years), bin 9 (41–45 years). Panel (a) shows the full sample. Panel (b) excludes workers younger than 61 years old and establishments established before 1985.

Second, we show that our results are not driven by the age of establishments in the sample and/or the retirement age of workers in Japan.[13] Establishment age is an important confounding factor, both because smaller establishments tend to be younger, so their workers naturally have shorter tenure and because younger establishments are more likely to employ younger workers (Ouimet and Zarutskie, 2014). In Japan, substantial worker churn occurs after 60, when the lifetime employment system ends with retirement. To account for the influence of establishment age and worker retirement on our results, we recalculate average worker experience/tenure in Panel (b) of Table 2 by restricting our sample to workers younger than 61 and establishments established before 1985 (i.e., excluding young establishments). We also remake the experience-tenure profiles using the same restricted sample in Panel (b) of Figure 1. The documented tenure–experience patterns continue to hold in this restricted sample, which we use to calculate the moments used to calibrate the model.

We conduct three additional checks to substantiate that our empirical findings are not driven by establishment age and/or workers' retirement age. One, regression results in Appendix Table B.5 show that larger establishments employ younger workers with longer tenure even after controlling for establishment age and various sets of fixed effects.[14] Two, the tenure–experience patterns hold within establishment-age bins across different sizes (see Appendix Figures A.4). Three, the differential experience/age–tenure profiles across establishments persist within these establishment

[13] The mandatory retirement age was 60 in the 1990s and the 2000s.

[14] In this regression, we also control for the average education of workers at the establishment level. As the average education of workers is strongly and negatively correlated with average worker age/experience, it is also negatively correlated with average worker tenure.

age bins (see Appendix Figure A.5).

Finally, the patterns we observe are consistent at the firm level. Our baseline results document heterogeneity across establishment size. Given that modern economies are dominated by multi-establishment firms, we also examine whether these patterns hold at the *firm* level. In particular, we ask whether differences in average tenure and experience are driven by size heterogeneity across firms or across establishments within firms.[15] Using firm-level employment data (grouped as 5–9, 10–29, ..., above 5,000), we group establishments by their parent firms and replicate the analysis, excluding young establishments and limiting workers to age 60 or younger for consistency with previous robustness checks. Appendix Table B.6 shows that larger firms employ younger workers with longer tenure on average, and Figure A.6 confirms that the steeper experience–tenure profile persists at the firm level. Overall, these stylized facts are driven by size heterogeneity across *firms*.

The empirical findings suggest that large and small establishments follow different hiring and firing strategies. Large (productive) establishments hire disproportionately more young workers, while small (less productive) establishments hire fewer. As workers age, some leave large establishments. Departing experienced workers from large establishments are often hired by small establishments, and most workers in small establishments remain there. In short, large establishments *selectively retain* workers over their life cycle, whereas small establishments adopt a *gradual* hiring strategy, bringing in more middle-aged and older workers. This mechanism helps explain why the experience tenure profile is steeper in large establishments.

3 Model: Efficient Worker Composition

This section presents a framework for understanding the efficient composition of workers and their allocation across establishments. We begin with a static model to analyze how shifts in the labor supply composition across worker groups affect labor allocations and aggregate output. We then extend the framework to a dynamic setting, incorporating the costs of reallocating labor across establishments, and examine how labor market frictions influence both allocations and efficiency.

3.1 Static Framework

We consider an economy with a continuum of heterogeneous establishments, each characterized by productivity $z \sim \Phi(z)$. Each establishment produces a unique consumption good by employing a *portfolio* of labor inputs drawn from different *types* of workers. At any point in time, each worker

[15] Tenure information in the dataset is reported at the firm level, making it natural to aggregate establishments to firms to study how employer size affects average tenure.

belongs to a group $g \in \mathcal{G}$, defined by characteristics such as skill, experience, and other productivity-relevant attributes. Markets clear at the g-level, and the set of worker groups is finite ($|\mathcal{G}| = G < \infty$), with all characteristics fully observable to all agents in the economy.

An establishment with productivity z takes the wage schedule $w_{g\,g\in\mathcal{G}}$ as given and solves:

$$\max_{l_g \geq 0} F\left(\{l_g\}; z\right) - \sum_g w_g l_g,$$

where $\{l_g\} = (l_1, ..., l_G)$ denotes the set of labor allocations. The first-order conditions are:

$$[l_g(z)]: \ \frac{\partial F(\{l_g\}; z)}{\partial l_g} \leq w_g, \quad \forall z, g. \tag{1}$$

The existence of a unique solution is guaranteed if the production function $F(\cdot)$ is increasing and jointly concave in l_g. Let $l_g^c(z)$ denote the corresponding optimal labor demand of group g for a firm with productivity z.

The economy is populated by a total mass L of workers, each endowed with one unit of labor supplied inelastically. Let L_g denote the supply of type-g workers. All G labor markets are perfectly competitive, with both firms and workers taking wages w_g as given. The labor market clearing condition for each group g is:

$$\int l_g^c(z) \ d\Phi(z) = L_g \quad \forall g. \tag{2}$$

Competitive Equilibrium A competitive equilibrium is characterized by a set of labor allocations $\{l_g^c(z)\}$ for each firm z and type of worker g, and a set of wages w_g such that firms optimize (1) and markets clear (2).

Planner's Problem: Allocation of Workers Across Groups

We now consider a problem in which the planner can reallocate labor supply across worker groups while keeping the total amount of labor fixed.[16] This setup captures the potential effects of policies that alter the age or skill structure of the workforce, such as those affecting fertility, retirement, immigration, or skill formation.

To characterize outcomes under varying degrees of flexibility in reallocating labor across groups, let $\mathcal{P}$ be a partition of $\mathcal{G}$, where each element $P_s \in \mathcal{P}$ is a set of groups g whose labor supplies the

[16] If the production function exhibits constant returns to scale in all labor inputs, we can normalize by total labor, focusing the analysis on the shares of each worker group in the total population.

planner can adjust jointly. If $g, g' \in P_s$, the planner can freely reallocate labor between these two demographic groups. Let L_s denote the total labor in partition element P_s, with the aggregate constraint $\sum_s L_s = L$.

Taking $\mathcal{P}$ as given, the planner's problem is,

$$\max_{l_g(z)\geq 0} \int F\left(\{l_g\}; z\right) \phi(z) dz \quad s.t \int \left(\sum_{g \in P_s} l_g(z) \right) \phi(z) dz = L_s, \ \forall s.$$

The associated Lagrangian is,

$$\mathcal{L} = \int F\left(\{l_g\}; z\right) \phi(z) dz + \sum_s \mu_s \left(L_s - \int \left(\sum_{g \in P_s} l_g(z) \right) \phi(z) dz \right).$$

The first-order conditions imply that the marginal product of labor is equalized within each labor-adjustable partition,

$$[l_g(z)] : \ \frac{\partial F\left(\{l_g\}; z\right)}{\partial l_g} \leq \mu_s \quad \forall z, g \in P_s, \forall s \tag{3}$$

Recall that $l_g^c(z)$ denotes the competitive labor allocation, and let $l_g^p(z)$ denote the planner's optimal labor allocation, with Y^c and Y^p representing the corresponding total outputs. Define aggregate labor allocations as $L_g^c \equiv \int l_g^c(z) d\Phi(z) = L_g$ and $L_g^p \equiv \int l_g^p(z) d\Phi(z)$. The planner's problem yields positive output gains from labor reallocation whenever the partition $\mathcal{P}$ is not composed entirely of singleton sets g for $g \in \mathcal{G}$. A straightforward feasibility argument establishes the following result (all proofs are provided in Appendix C.1).

Proposition 1. *If the number of elements in the partition $\mathcal{P}$ is smaller than the total number of groups ($S < G$), then the planner's output is weakly larger than in the competitive equilibrium, i.e., $Y^p \geq Y^c$. The inequality is strict if there exist $g, g' \in P_s$ for some s such that $w_g \neq w_{g'}$.*

Moreover, if two demographic groups belong to the same partition element and face different wages, the planner optimally allocates relatively more aggregate labor to the group with the higher wage.

Proposition 2. *If $g, g' \in P_s$ with $w_g > w_{g'}$, then $\frac{L_g^p}{L_{g'}^p} > \frac{L_g^c}{L_{g'}^c} = \frac{L_g}{L_{g'}}$.*

Intuitively, higher wages in a competitive market reflect higher marginal productivity. Because the planner optimally equalizes marginal productivity across groups within a partition element, the aggregate labor supply of the group with higher marginal productivity must increase. Moreover, we can rank the potential gains from labor supply reallocation.

Definition 1. *Let $\mathcal{P}$ and $\mathcal{P}'$ be partitions of $\mathcal{G}$. We say that $\mathcal{P}$ is a refinement of $\mathcal{P}'$ if every element of $\mathcal{P}$ is a subset of some element of $\mathcal{P}'$. In this case, $\mathcal{P}$ is referred to as finer than $\mathcal{P}'$, and we write $\mathcal{P} \leq \mathcal{P}'$.*

Note that the finer-than relationship on the set of partitions of $\mathcal{G}$ defines a partial order. Let $Y^p(\mathcal{P})$ denote the total output associated with partition $\mathcal{P}$. A straightforward revealed-preference argument shows that refining a partition cannot reduce total output:

Proposition 3. *If $\mathcal{P}$ is a refinement of $\mathcal{P}'$, then $Y^p(\mathcal{P}') \geq Y^p(\mathcal{P})$.*

In particular, this implies that the maximal partition, i.e., that partition consisting of the entire set $\mathcal{P}^{max} = \{\mathcal{G}\}$ yields the largest potential output gains from labor reallocation. Conversely, the minimal partition, consisting of the singletons of $\mathcal{G}$, $\mathcal{P}^{min} = \{\{g\} : g \in \mathcal{G}\}$, corresponds to the competitive equilibrium.

Proposition 3 implies that total output is monotone with respect to the partial order on partitions,

Corollary 1. *$Y^p(\mathcal{P}^{max}) \geq Y^p(\mathcal{P})$ for any partition $\mathcal{P}$, with strict inequality if there exist groups g, g' with $w_g \neq w_{g'}$ that belong to different elements of $\mathcal{P}$. Moreover, $Y^p\left(\mathcal{P}^{min}\right) = Y^c$.*

CES Production Technology

To proceed, we assume that the production technology aggregates worker groups using a CES function with decreasing returns:

$$Y = \int y(z)d\Phi(z) = \int \left(\sum_g A_g(z)^{\frac{1}{\sigma}} l_g(z)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\alpha\sigma}{\sigma-1}} d\Phi(z)$$

where $y(z)$ denotes output of firm type z, $A_g(z)$ is the match-specific productivity for group g at firm z, $\sigma \in (0, \infty)$ governs substitutability across worker groups, and $\alpha \in (0, 1)$ determines returns to scale. This specification captures two empirically grounded features: decreasing returns at the firm level and imperfect substitutability across worker types.[17]

With this functional form, profit maximization yields the following labor demand and wage deter-

[17] Decreasing returns are standard in firm dynamics models following Hopenhayn (1992) and Melitz (2003). Imperfect substitutability is supported by empirical evidence in Card and Lemieux (2001), Borjas (2003), and Ottaviano and Peri (2012).

mination:

$$l_g(z) = \frac{y(z)^{\sigma-\frac{\sigma-1}{\alpha}} A_g(z)}{\int y(z')^{\sigma-\frac{\sigma-1}{\alpha}} A_g(z') d\Phi(z')} L_g \tag{4}$$

$$w_g = \alpha \left(\frac{\int y(z)^{\sigma-\frac{\sigma-1}{\alpha}} A_g(z) d\Phi(z)}{L_g} \right)^{1/\sigma} \tag{5}$$

Equation (4) shows that equilibrium employment of group g at firm z depends on three factors: the group's total supply L_g, the firm's output-weighted productivity for that group relative to the market average, and the firm's scale. The wage equation (5) indicates that wages rise with a group's average productivity and fall with its supply, capturing both productivity and scarcity effects.

Substituting the labor demand functions into the production function gives an implicit characterization of equilibrium output:

$$y(z) = \left(\sum_g \frac{A_g(z) L_g^{\frac{\sigma-1}{\sigma}}}{\left(\int y(z')^{\sigma-\frac{\sigma-1}{\alpha}} A_g(z') d\Phi(z') \right)^{\frac{\sigma-1}{\sigma}}} \right)^{\frac{\alpha}{(\sigma-1)(1-\alpha)}} \qquad \forall z$$

Under the CES specification, Proposition 2 can be strengthened: the labor ratio inequality holds not only in aggregate but at every firm.

Proposition 4. *Under CES production, if $g, g' \in P_s$ with $w_g > w_{g'}$, then $\frac{l_g^p(z)}{l_{g'}^p(z)} > \frac{l_g^c(z)}{l_{g'}^c(z)}$ for all z.*

This stronger result follows from the CES property that the marginal rate of technical substitution between any two worker types depends only on their labor ratio and productivity parameters, ensuring that the direction of reallocation is uniform across firms.

Proceeding with the planner's problem under CES, the social planner maximizes aggregate output by choosing labor allocations subject to the economy-wide resource constraint:

$$\max_{l_g(z) \geq 0} \int \left(\sum_g A_g(z)^{\frac{1}{\sigma}} l_g(z)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\alpha\sigma}{\sigma-1}} d\Phi(z) \quad \text{s.t.} \quad \int \sum_g l_g(z) d\Phi(z) = L$$

The first-order conditions equalize the marginal product of labor across all worker-firm pairs:

$$\alpha A_g(z)^{\frac{1}{\sigma}} l_g(z)^{-\frac{1}{\sigma}} y(z)^{1-\frac{\sigma-1}{\alpha\sigma}} = \lambda \quad \forall g, z$$

where λ is the shadow value of labor. This yields the optimal allocation:

$$l_g^*(z) = \frac{A_g(z)y^*(z)^{\sigma-\frac{\sigma-1}{\alpha}}}{\sum_{g'}\int y^*(z')^{\sigma-\frac{\sigma-1}{\alpha}}A_{g'}(z')d\Phi(z')}L \tag{6}$$

Comparing equations (4) and (6) highlights the key difference: in competitive equilibrium, each group's labor supply L_g is allocated across firms separately, whereas the planner allocates the total labor supply L jointly across firms and all groups. This flexibility enables the planner to exploit productivity differences across firms and workers more efficiently.

The optimal output levels satisfy:

$$y^*(z) = \left(\frac{\sum_g A_g(z)L^{\frac{\sigma-1}{\sigma}}}{\left(\sum_{g'}\int y^*(z')^{\sigma-\frac{\sigma-1}{\alpha}}A_{g'}(z')d\Phi(z')\right)^{\frac{\sigma-1}{\sigma}}}\right)^{\frac{\alpha}{(\sigma-1)(1-\alpha)}}$$

As shown in Proposition 3, aggregate output under the optimal allocation exceeds that in competitive equilibrium whenever groups are imperfect substitutes ($\sigma < \infty$). The size of this gain depends on the dispersion of productivity ratios $A_g(z)/A_{g'}(z)$ across firms—greater heterogeneity in comparative advantage generates larger potential gains from reallocation.[18]

This framework provides the theoretical foundation for analyzing efficiency tradeoffs in labor recomposition across groups. In empirical and quantitative applications, however, it is necessary to account for dynamic worker characteristics. As workers transition between groups, the composition of the labor force evolves, influencing agents' equilibrium choices. These dynamics constrain the planner's feasible policy set.

3.2 Dynamic Model

In this section, we extend the model to examine how dynamic considerations influence optimal labor allocation. This is important as the composition of groups can evolve over time. We consider two dimensions along which worker groups differ: experience and skill. Experience evolves deterministically as workers age, while skill captures the quality of a worker's match to their job and can evolve stochastically over time. Firms in our model do not differ in industry characteristics, so skill in the model does not capture occupation or industry characteristics or the worker's education, but rather reflects the worker's ability in their particular job, which is transferable across firms. The model includes a reallocation friction in the form of a firing cost to illustrate the impact of these

[18] The dispersion of A_g (integrated across firm types) across worker groups also matters for the aggregate gain, as workers from different groups are imperfect substitutes.

dynamic factors on the optimal labor composition. While firing costs are empirically relevant in the Japanese context, the insights of the dynamic model also apply to other types of frictions.

In the model, heterogeneous firms employ a portfolio of workers in different experience and skill groups, and this portfolio evolves within firms over time. The model allows us to characterize these optimal portfolios and the hiring strategies that support them.

Workers The model economy consists of overlapping generations of workers characterized by experience $X := (x_1, x_2, ..., x_J)$. Workers with the highest experience level, x_J, retire and are replaced by a new cohort entering at experience $j = 1$. Workers also face random labor-market participation shocks that capture unmodeled entry and exit, such as death or illness, marriage, or childbirth. Hence, the mass of workers in each experience bin is denoted by μ_j, which determines the size of each cohort relative to the mass of entrants, normalized to $\mu_1 = 1$. In our quantitative analysis, we allow for $\mu_{j'} > \mu_j$ for $j' > j$. This could arise if there are fluctuations in labor force participation by age.

Throughout their careers, workers accumulate both experience and skill that is valuable to all firms.We assume that work experience is acquired *deterministically*, while skill evolves *stochastically*. Specifically, a worker's experience evolves with age, so older workers are necessarily more experienced than younger ones. By contrast, skill evolves through stochastic transitions. We represent skill as $h \in H := \{h_1, h_2, ..., h_I\}$. Let Ψ_j be the $I \times I$ transition matrix, where $\psi^j_{i,i'}$ is the probability of moving from skill level h_i to $h_{i'}$ for an individual with experience x_j. The probablity of skill transitions depend solely on experience and are independent of the employing firm. At birth, workers' skill is distributed according to $g_0(h_i)$, the measure of individuals with level h_i upon entering the economy.

Firms Each period, firms hire a portfolio of workers defined by their skill h_i and experience level x_j. Let L_{ij} (a firm-level state variable) denote the number of workers in the (i, j) bin at the beginning of the period, and l_{ij} (a firm-level choice variable) the number hired from this bin during the period. The model features two types of job separations. First, the firm may choose to dismiss workers, incurring a firing cost equal to a fraction $\tau > 0$ of the worker's wage in that period. Second, a randomly selected fraction $1 - \eta$ of incumbent workers separate from the firm for exogenous reasons, such as job-to-job transitions.

The presence of firing costs introduces a non-trivial dynamic trade-off: firms must weigh the benefits of moving toward their frictionless optimal employment level against the costs of adjustment. This creates an *inaction region*, a range of employment levels where the firm optimally maintains its current workforce. Two factors influence the characterization of these regions. First, skill

evolves stochastically, creating uncertainty about future workforce composition regardless of current choices. Second, worker groups are interdependent: employment decisions for one group alter the firm's future state and influence optimal choices for other groups via the production function. We now formalize this problem and derive the optimal policy.

A firm with productivity z chooses labor l_{ij} to maximize the value function,

$$V(z, \{L_{ij}\}) = \max_{\{l_{ij}\}} \sum_{ij} F\left(\{l_{ij}\}, z\right) - \sum_{ij} w_{ij} l_{ij} - \sum_{ij} \eta\tau w_{ij} (L_{ij} - l_{ij})^{+} + \beta \sum_{j=1}^{J-1} \sum_{i,i'} \mu_j \psi^j_{i,i'} V(z, \{\tilde{L}_{i'j+1}\}). \tag{7}$$

The first-order conditions (FOCs) of the optimization problem (7) are:

$$[l_{ij}]: \ \frac{\partial F}{\partial l_{ij}} = w_{ij} \left(1 - \eta\tau \mathbb{1}\left[L_{ij} > l_{ij}\right]\right) - \beta\mu_j \sum_{i'} \psi^j_{ii'} \frac{\partial V(z, \tilde{L}_{i'j+1})}{\partial L_{i'j+1}} \quad \forall i, j.$$

The Envelope Condition in the current and next periods implies that

$$\frac{\partial V}{\partial L_{ij}} = -\eta\tau w_{ij} \mathbb{1}\left[L_{ij} > l_{ij}\right].$$

Together, FOCs imply

$$\frac{\partial F}{\partial l_{ij}} = w_{ij} \left(1 - \eta\tau \mathbb{1}\left[L_{ij} > l_{ij}\right]\right) + \beta\eta\tau\mu_j \sum_{i'} \psi^j_{ii'} w_{i'j+1} \mathbb{E}\left[\mathbb{1}\left[\tilde{L}_{i'j+1} > l_{i'j+1}\right]\right]. \tag{8}$$

Comparing (8) with (1) reveals two additional effects of hiring when a firing cost is present. First, if the current employment l_{ij} is below the existing stock L_{ij}, the firm saves on firing costs. Second, if next period's optimal employment $l_{i,j+1}$ is below the projected stock $\tilde{L}_{i,j+1}$, the firm faces a potential firing cost.

We now derive the firm's optimal hiring and firing strategies, which imply an inaction region for employment given the current stock L_{ij}. Let FC_{ij} denote the expected continuation cost associated with choosing l_{ij} workers,

$$FC_{ij} := \eta\tau\mu_j \sum_{i'} \psi^j_{ii'} w_{i'j+1} \mathbb{E}\left[\mathbb{1}\left[\tilde{L}_{i'j+1} > l_{i'j+1}\right]\right].$$

If the firm hires additional workers ($l_{ij} \geq L_{ij}$), the first-order condition for the (i, j) worker group

becomes

$$\frac{\partial F}{\partial l_{ij}} = w_{ij} + \beta FC_{ij}, \tag{9}$$

which yields a unique solution for l_{ij}, denoted by l_{ij}^{lb}. Next, if the firm fires workers ($l_{ij} < L_{ij}$), the first-order condition is

$$\frac{\partial F}{\partial l_{ij}} = w_{ij}(1 - \eta\tau) + \beta FC_{ij} \tag{10}$$

which yields a unique value l_{ij}^{ub}. Since $l_{ij}^{lb} < l_{ij}^{ub}$, the inaction region is defined by these bounds; if $L_{ij} \in [l_{ij}^{lb}, l_{ij}^{ub}]$, the firm chooses neither to hire nor fire workers.

The firm's labor choice, governed by equations (9) and (10), depends on employment decisions from both the previous and the next period. Moreover, the choice of l_{ij} depends not only on the current wage w_{ij} but also on future wages for the same and adjacent worker groups, since these influence expected firing costs. Focusing on the steady state, where wages are constant over time, the firm only needs to form correct expectations of the current period's prevailing wages and the expected future firing probabilities.

Labor Markets Labor markets in this economy are indexed by (i, j), where i denotes the skill bin h_i and j the experience bin x_j. In equilibrium, labor demand must equal labor supply in each (i, j) submarket:

$$\int_{z_{min}}^{\infty} l_{ij}(z) d\Phi(z) = g(h_i, x_j); \quad \forall\ (i, j), \tag{11}$$

where $l_{ij}(z)$ denotes the number of workers with skill h_i and experience x_j employed by an establishment with productivity draw z, and $g(h_i, x_j)$ is the labor supply of workers with skill h_i and experience x_j. In addition, Equation (11) determines the market-specific wage w_{ij}.

In Appendix C.6, we develop an algorithm to solve for the equilibrium. The key idea is to determine the "correct" expectations of employment choices for the next experience group, since these expectations influence current employment decisions and, consequently, the equilibrium in each labor submarket. In equilibrium, these expectations must be consistent with the actual employment choices firms make for the next experience group.

4 Model Calibration and Fit

In this section, we apply our model to Japanese worker–establishment data. We begin by describing the calibration procedure and the moments we target. We use the static model with no firing costs

as our baseline case for calibration and then study how dynamic features influence the optimal composition of workers across firms.[19]

4.1 Calibration Specification

For our quantitative exercises, we parameterize both the production function and the stochastic learning process. We assume there are K types of firms that differ in productivity. Each firm of type $k \in \{1, ..., K\}$ operates a CES production function with decreasing returns to scale given by:

$$F(\{l_{ij}, z_k\}) = \left(\sum_{ij} A_{ijk}^{\frac{1}{\sigma}} l_{ijk}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\alpha\sigma}{\sigma-1}},$$

where l_{ijk} denotes the labor input from worker type (i, j) employed by a type-k firm. The parameter $\sigma \in (0, \infty)$ represents the elasticity of substitution across worker types, capturing imperfect substitutability in production. The parameter $\alpha \in (0, 1)$ governs the degree of decreasing returns to scale. This specification also incorporates firm-worker-specific productivity terms, A_{ijk}, that capture match-specific complementarities between worker and firm types. Identifying these productivity parameters from observed data moments is a central component of our quantitative analysis.

We specify the skill transition process Ψ^j as a learning–depreciation chain, in which workers can move up or down the skill ladder by only one step in each period. This parsimonious formulation captures two key forces: skill acquisition through experience-dependent learning (with probability p_j) and constant skill depreciation due to knowledge obsolescence (with probability q).

For workers with skill h_i where $1 < i < I$, the transition probabilities are:

- Skill improvement: $\psi^j_{i,i+1} = \Pr(h_{i+1}|h_i, x_j) = p_j(1-q)$ — learning occurs and knowledge does not depreciate.
- Skill depreciation: $\psi^j_{i,i-1} = \Pr(h_{i-1}|h_i, x_j) = q(1-p_j)$ — knowledge depreciates without learning.
- Skill maintenance: $\psi^j_{i,i} = \Pr(h_i|h_i, x_j) = 1 - p_j - q + 2p_jq$ — either both events occur (canceling out) or neither happens.

At the boundaries, we impose reflecting conditions. Workers at the lowest skill level ($i = 1$) cannot depreciate further, so $\psi^j_{1,0} = 0$, $\psi^j_{1,1} = 1 - p_j$, and $\psi^j_{1,2} = p_j$; similarly, workers at the highest level ($i = I$) cannot advance, so $\psi^j_{I,I+1} = 0$, $\psi^j_{I,I} = 1 - q$, and $\psi^j_{I,I-1} = q$.

[19] In Section 5.3, we recalibrate a set of (important) model parameters in a setting with a positive firing cost.

4.2 Target Moments and Estimation

We estimate the model using Japanese establishment data from 2005–2011. Workers are classified into $J = 9$ experience bins: $x_j \in \{1\text{-}5, 6\text{-}10, ..., 36\text{-}40, 40+\}$ years, and establishments are divided into two size categories: small (fewer than 50 employees) and large (50 or more employees), indexed by $k \in \{\mathrm{S}, \mathrm{L}\}$. Skill is unobserved and takes two values, $h_i \in \{\mathrm{L}, \mathrm{H}\}$, representing low-skill and high-skill workers, respectively. All wages are normalized by the economy-wide average wage.

We first set the values of some parameters externally from our calibration approach. We set the discount factor to $\beta = 0.815$, consistent with a five-year period and an annual discount rate of 0.96. The elasticity of substitution is set to $\sigma = 2.5$, implying moderate substitutability across worker types.[20] Following Hopenhayn et al. (2022) and Karahan et al. (2024), we set the returns-to-scale parameter to $\alpha = 0.6$, reflecting the labor share in production. The initial skill distribution, $g_0 = [0.4262, 0.5738]$, for low- and high-skill workers matches the share of college-educated workers entering the labor force during 2005–2011. For the baseline model, we set firing costs to $\tau = 0$, though we assume an infinitesimally small firing cost to measure job separations of different types from the model. The assumption of no firing costs provides an important benchmark in the static model. Section 5.3 assesses the sensitivity of our results to these calibrated parameters.

After setting the values for external calibration we have 46 remaining parameters to estimate: 36 match-specific productivity terms A_{ijk} (for $i \in \{L, H\}$, $j \in \{1, ..., 9\}$, $k \in \{S, L\}$), 8 experience-specific learning rates p_j for $j \in \{1, ..., 8\}$ (with $p_9 = 0$ for the most experienced workers), one skill depreciation rate q, and the exogenous separation rate η. To ensure parsimony and interpretability, we restrict the learning probabilities p_j to follow a unimodal pattern over the experience distribution. These parameters are identified through distinct sets of moments: employment shares and wage profiles primarily pin down the static match productivities, A_{ijk}, while tenure-experience profiles and the tenure distribution capture the dynamic learning rates, p_j, skill depreciation, q, and exogenous separation rate, η, as we formalize in the identification discussion below.

Since worker experience and firm employment are observable in the data, we target the following five sets of moments. First, we target employment shares l_{jk} in aggregate labor demand, that is, the fraction of type-j workers employed by type-k firms as a share of the total number of workers. Second, we target the mean wages of type-j workers in type-k firms, w_{jk}. Third, we target tenure-experience profiles defined as the average tenure by experience bin-j within each firm type-k. Fourth, we target the tenure distribution measured as the distribution of employment across firm type-k for each worker tenure bin-t. Finally, we target the average firing rate measured

[20] Papers such as Card and Lemieux (2001), Borjas (2003), and Ottaviano and Peri (2012) estimate that the elasticity of substitution across workers in different experience groups is between 2 and 5. The elasticity of substitution across workers with different skill levels is estimated to be between 1.2 and 1.5 in Katz and Murphy (1992) and Card and Lemieux (2001).

Table 3: Parameters

Parameter	Description	Value
Externally Calibrated		
β	Discount rate (4 years)	0.815
σ	Elasticity of substitution inputs	2.5
α	Elasticity of substitution inputs	0.6
g_0	Initial distribution of skills	[0.4262, 0.5738]
τ	Firing costs	$\epsilon \to 0$
Simulated Method of Moments		
A_{ijk}	Worker-establishment productivities	See Figure 3
p_j	Learning probability	[0.7756, 0.0195, 0.0195, 0, 0, 0, 0, 0]
q	Knowledge depreciation/obsolescence	0.0879
η	Exogenous separation	0.8882

Note: Table lists the externally calibrated parameters and the ones estimated using the Simulated Method of Moments.

as the economy-wide rate of involuntary job separations.

These moments yield 64 targets for estimation. Specifically, we have $J \times K = 18$ employment shares, average wages, and tenure-experience observations, $J = 9$ tenure distribution moments, and the aggregate firing rate.[21]

To ensure economic interpretability, we impose the constraint that wages increase monotonically in skill:

$$w_{H,j} > w_{L,j} \quad \forall j \in \{1, ..., 9\} \tag{12}$$

where $w_{i,j}$ denotes the model-implied wage for skill type i and experience j. This constraint ensures that high-skill workers earn strictly more than low-skill workers with identical experience and is justified on the grounds that high-skill workers could also do low-skill jobs.

Collectively denoted by Θ, we estimate the parameters by minimizing the weighted sum of squared deviations:

$$\hat{\Theta} = \arg\min_{\Theta} \sum_{m} \omega_m \frac{(d_m - m_m(\Theta))^2}{d_m^2 + m_m(\Theta)^2} \tag{13}$$

[21] While the tenure distribution could in principle generate 9×2 moments, aggregation across firm types for each experience level eliminates redundancy, yielding 9 independent moments. Precisely speaking, there is one redundant condition in the first set of moments, as the employment shares across worker experience bins and firm types sum to one.

Table 4: Moments

Moment	Data	Model
1. Employment-Experience Distribution	See Figure 2a	See Figure 2a
2. Wage-Experience Profile (normalized)	See Figure 2b	See Figure 2b
3. Tenure-Experience Profile	See Figure 2c	See Figure 2c
4. Employment-Tenure Distribution	See Figure 2d	See Figure 2d
5. Firing Rate	2.2%	2.20%

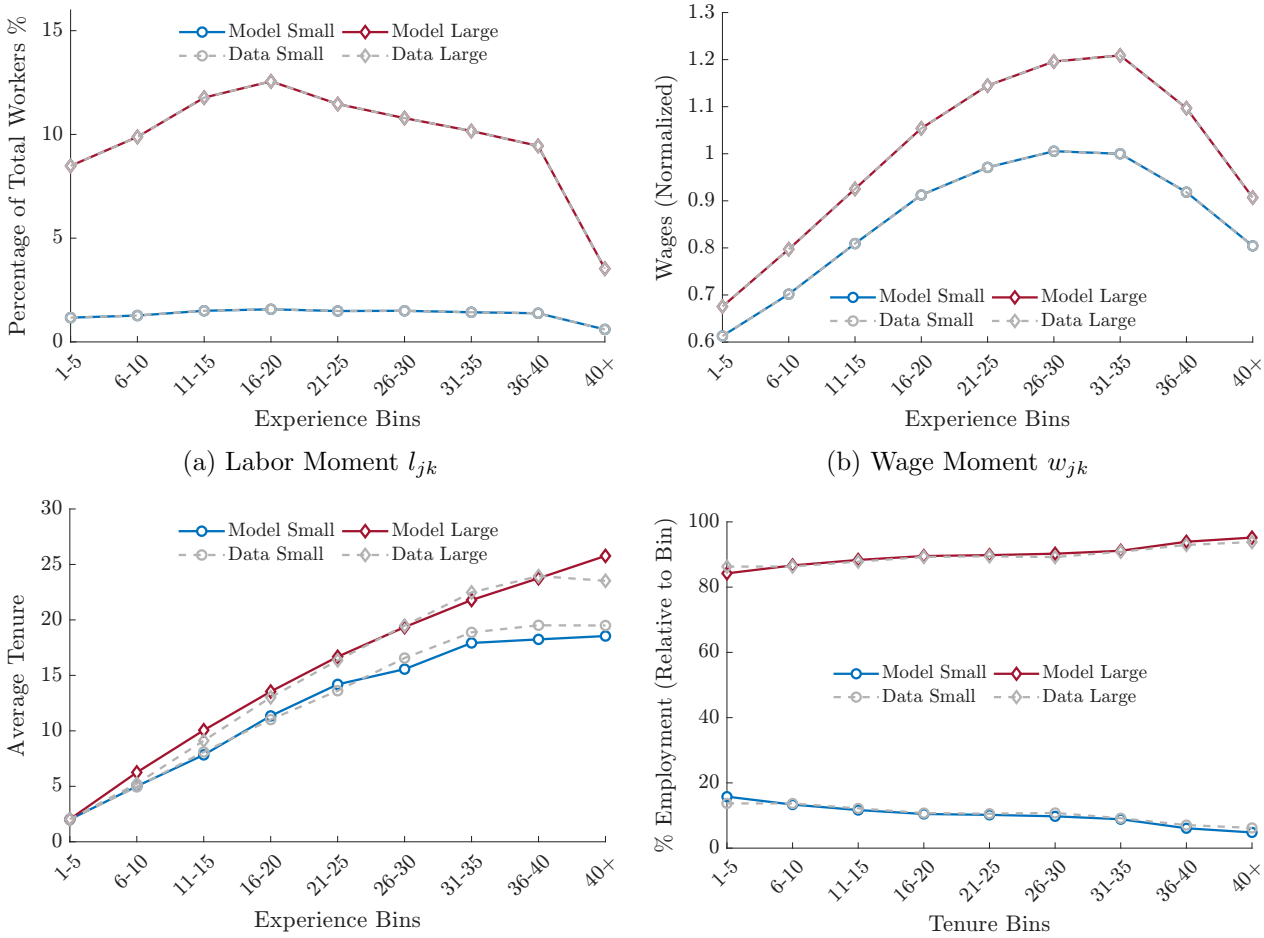
Note: Table lists the model-implied and data-targeted moments. Data moments are computed for the period 2005–2011.

where d_m and $m_m(\Theta)$ denote data and model moments, respectively. The denominator normalization prevents moments of different scales from dominating the objective function. The weights ω_m ensure equal importance across the five moment categories, with each category receiving a weight of 1/5. For example, the 18 employment share moments each receive weight $\omega_m = 1/(5 \times 18)$, while the single aggregate separation rate receives weight $\omega_m = 1/5$. The values of parameters resulting from the calibration are presented in Table 3.

Moments and Identification Table 4 summarizes the model's fit to the targeted moments, and Figure 2 compares model predictions with the data. The model reproduces key features of the Japanese labor market, including heterogeneous tenure–experience profiles across firm types and the aggregate firing rate. The identification strategy uses distinct moment subsets to recover different parameter groups. The static productivity parameters A_{ijk} are identified from the employment shares by experience and firm type (Figure 2a) and the wage–experience profiles (Figure 2b). Combined with the monotonicity constraint on wages across skill levels, these moments uniquely determine the 36 match-specific productivities. In Appendix C.5 we formalize this argument.

The dynamic parameters—the exogenous separation rate η, experience-specific learning rates p_j, and skill depreciation rate q—are identified from the remaining moments. The tenure–experience profiles by firm type (Figure 2c) capture how job stability varies with worker experience and firm types. The overall tenure distribution (Figure 2d) and aggregate firing rate provide variation to disentangle learning dynamics from exogenous job destruction. This strategy exploits the fact that endogenous separations depend on the evolution of match-specific productivity through skill accumulation, whereas exogenous separations affect all matches uniformly.

Figure 2: Moments Fit



(a) Labor Moment l_{jk}

(b) Wage Moment w_{jk}

(c) Experience-Tenure

(d) Tenure Distribution

Notes: This figure displays the model fit for the first four sets of targeted moments. Panels (a) and (b) show employment shares and average wages by experience group and firm type. Panel (c) presents average tenure by experience within each firm type. Panel (d) shows the employment shares of large/small firms (i.e., firm type) for workers in each tenure bin (i.e., 1-5, 6-10, ...,40+). Grey dashed lines represent data moments; solid lines represent model predictions. Small firms have fewer than 50 employees; large firms have 50 or more employees.

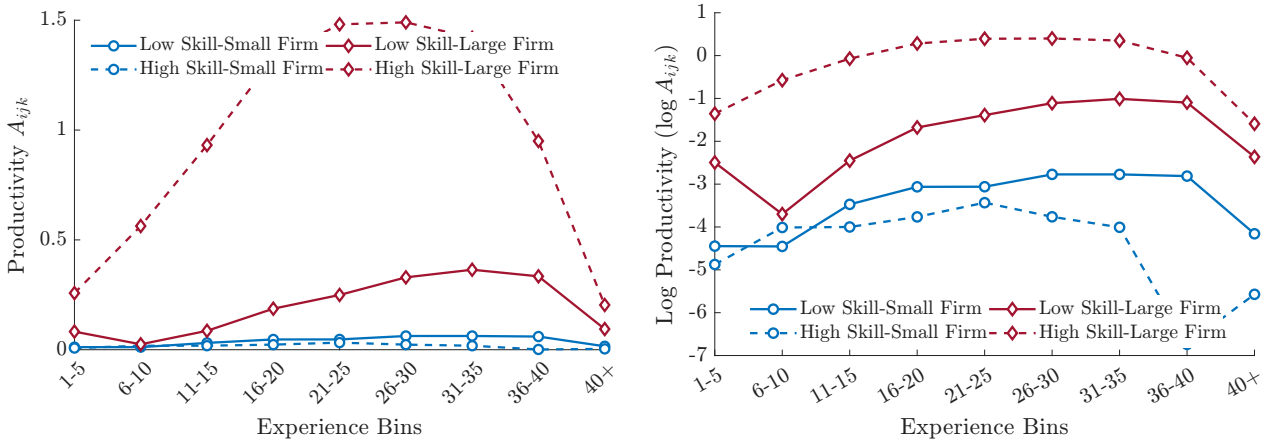
4.3 Estimation Results and Hiring Dynamics

Log Supermodularity in Productivity Parameters

Figure 3 shows the estimated match-specific productivities A_{ijk}, a key output of the structural estimation. They are identified non-parametrically from observed employment distributions and wage differentials across firm types.

The results reveal two patterns in worker–firm complementarities. First, productivity levels are

Figure 3: Estimated Worker-Firm Productivities across Experience A_{ijk}



(a) Productivity A_{ijk} (b) Log Productivity $\log A_{ijk}$

Notes: Estimated match-specific productivities for skill type $i \in \{L, H\}$, experience group $j \in \{1, ..., 9\}$, and firm type $k \in \{S, L\}$. Panel (a) shows productivity levels; panel (b) shows log productivities. Each line represents a skill-firm pair, with productivity varying across experience groups.

higher in large firms than in small firms for both skill types, consistent with standard theories of firm heterogeneity. Second, in small firms, low-skill workers generally high higher productivity parameters than high-skill workers across experience levels. Combined with the productivity advantage of high-skill workers in large firms, this pattern implies strong skill–size complementarity or *log supermodularity* in aggregate production.

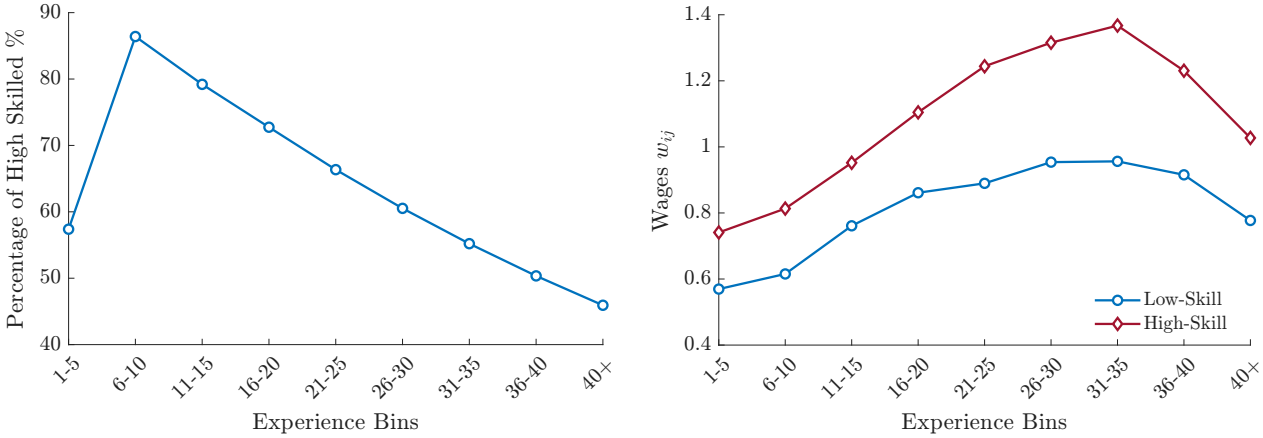
Skill and Wages

A key feature of our approach is that skill is unobserved and inferred from the data. Figure 4 presents the estimated skill composition and wage profiles implied by this identification strategy.

Figure 4a shows a pronounced life-cycle pattern in skill composition: the share of high-skill workers rises from 57% at labor market entry to 75% at 6–10 years of experience, then declines steadily to below 50% after 36+ years. This inverted-U reflects the interplay of learning and depreciation—early-career workers experience rapid skill accumulation (learning probabilities exceed 50% in the first experience bin; see Table 3), but the constant depreciation rate q eventually dominates, leading to net skill deterioration in later careers. Remember that the skill in the model should be interpreted as the worker's fit for their given task. These patterns imply a process in which young workers match with tasks early in their career, with the fit then slowly degrading over the life cycle.

Panel 4b shows that both skill types have hump-shaped wage profiles, with key differences. High-skill workers experience faster early-career wage growth, widening the skill wage gap from 15% at

Figure 4: Percentage of High-Skill and Wages $\alpha = 0.6$

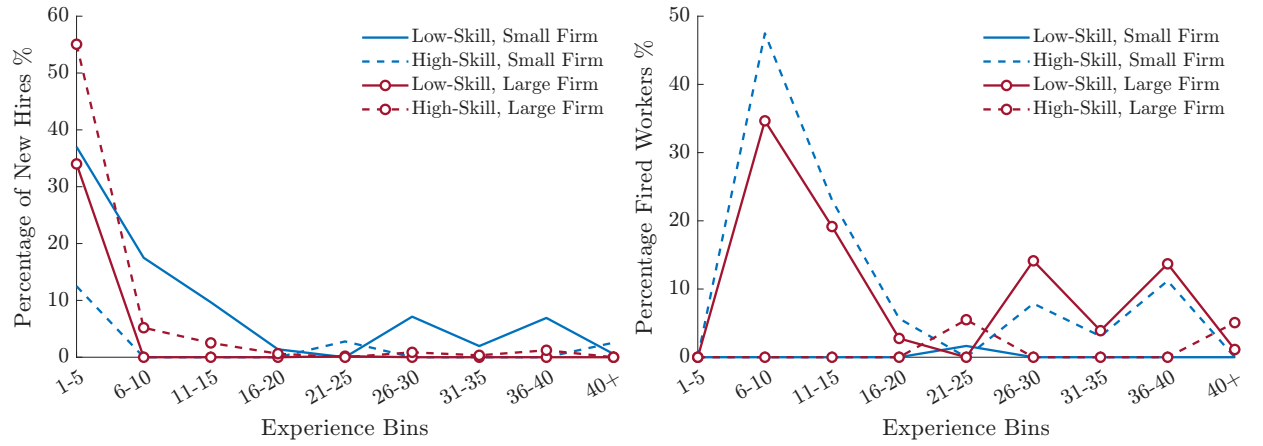


(a) Percentage of High Skill

(b) Wages by Skill

Notes: Panel (a) shows the share of high-skill workers by experience group. Panel (b) shows wage profiles by skill type across experience groups.

Figure 5: Hiring and Firing Strategies: Up-or-Out Dynamics in Large Firms



(a) Hiring

(b) Firing

Notes: This figure shows the percentage of new hires and fired workers by firm size and skill level for the case $\tau = 0$. Percentages are calculated relative to total hiring and firing within each firm.

entry to over 35% at peak earnings. This life-cycle increase in wage inequality arises endogenously from the interaction of skill dynamics and match-specific productivities.

Firms' Hiring and Firing Strategies

We use our calibrated model to study the hiring and firing strategies of establishments of different sizes, showing that large, high-productivity establishments disproportionately employ young workers. We then contrast these predictions with Japanese establishment-level data. Figure 5 highlights the patterns of hiring and firing by worker group across firms.[22] Panel (a) shows that nearly 90% of newly hired workers in large establishments are job market entrants (experience 1–5), compared with about 50% in small establishments. Small establishments hire a larger share of middle-aged and older workers and employ disproportionately more unskilled workers. Panel (b) shows that small establishments fire disproportionately more young, skilled workers, whereas large establishments tend to dismiss more experienced, unskilled workers. These differences in hiring and firing behavior stem from calibrated firm–worker group-specific productivities and result in heterogeneous exposure to changes in worker skills and experience across establishments.

Additionally, small and large firms separate from roughly the same number of workers (1.03% and 1.17% of the labor force, respectively). However, since small firms employ a much smaller share of the labor force (12%) and the exogenous separation rate is identical across firms, small firms exhibit a higher worker separation rate than large firms. This pattern is consistent with the key empirical finding of Papageorgiou (2018).

Next, to help validate the predictions of the model, we assess whether the hiring strategies predicted by our model, which posits that more productive establishments hire disproportionately more young workers, especially job-market entrants, are consistent with the patterns in the Japanese data. Our data include the number of new entrants to the labor market hired at the establishment-year level. For establishments observed in consecutive years, we can then infer net incumbent hiring as the difference between employment changes and entrant hires. We examine two questions: (1) do larger establishments hire disproportionately more entrants? and (2) are larger establishments more likely to hire entrants relative to incumbents? We use the dataset from the calibration exercise to address these questions.

Table 5 shows that larger establishments hire more job-market entrants overall (col. 1), disproportionately relative to total employment (col. 2), and relative to net hiring (col. 3). Table 6 further indicates that they are more likely to hire entrants than incumbents (or to reduce incumbent employment). Together, these findings support our model's predictions about heterogeneous hiring strategies across establishments.[23]

[22] Figure 5 does not include exogenous separations and the associated replacement hiring when calculating the number of fired and hired workers. Because the exogenous separation rate is constant across establishments and worker groups, including them would not change the relative patterns of hiring and firing of worker groups across firms.

[23] Data limitations prevent us from conducting additional tests. First, we lack information on the firing and hiring

Table 5: Hiring patterns by establishment size (2005–2011)

	(1) Log(New Hires)	(2) New Hires/ Employment	(3) New Hires/ Net Hires
Log(Employment)	0.532***	0.004***	0.126***
	(0.004)	(0.000)	(0.003)
Establishment Age	−0.001***	−0.000***	0.002***
	(0.000)	(0.000)	(0.000)
Constant	−1.297***	−0.002***	−0.112***
	(0.017)	(0.000)	(0.011)
City FE	Yes	Yes	Yes
Industry × Year FE	Yes	Yes	Yes
Prefecture × Year FE	Yes	Yes	Yes
Observations	61,961	279,254	32,760
R^2	0.457	0.137	0.284

Notes: The dependent variable is the log number of new hires from job market entrants (column 1), the ratio of new hires to total employment (column 2), or the ratio of new hires to net employment change (column 3). New hires refers to job market entrants hired by the establishment. Net hires is the change in establishment employment between consecutive years. In column (3), the ratio is undefined when new hires are zero and net hiring is non-positive. Standard errors in parentheses, clustered at the firm level. Column (3) has fewer observations because it requires establishments to appear in consecutive years. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

5 Results: Gains from Optimal Labor Supply Composition

This section quantifies the efficiency gains from optimal labor composition by solving a planner's problem that takes the estimated productivities A_{ijk} as given and optimally allocates workers across experience and skill bins to maximize aggregate output. This exercise provides an upper bound on potential productivity gains from policies that alter workforce composition, as policies that influence workforce composition, like fertility incentives, education expansion, and retirement age reforms, are unlikely to fully optimize workforce composition by both skill and age.

Following Corollary 1, we implement the finest possible partition, $\mathcal{P}^{max} = \mathcal{G}$, in which each demographic group forms its own labor market. This approach provides an upper bound on reallocation gains by giving the planner full flexibility in assigning workers to firms, subject only to aggregate labor supply constraints. The resulting efficiency gain captures the gap between the decentralized equilibrium and the first-best allocation, highlighting the direction of and potential scope for policy

of job market incumbents. Second, we are unable to include establishment fixed effects in the regressions because establishments appear on average only twice in our 8-year sample.

Table 6: Hiring new entrants versus experienced workers (2005–2011)

	(1)	(2)	(3)	(4)
	Hire Entrants − Hire Experienced			
Log(Employment)	0.126^{***}	0.122^{***}	0.125^{***}	0.122^{***}
	(0.003)	(0.003)	(0.003)	(0.003)
Establishment Age		0.002^{***}		0.002^{***}
		(0.000)		(0.000)
Constant	-0.480^{***}	-0.534^{***}	-0.480^{***}	-0.532^{***}
	(0.008)	(0.010)	(0.009)	(0.010)
City FE	Yes	Yes	Yes	Yes
Prefecture FE	Yes	Yes	No	No
Industry FE	Yes	Yes	No	No
Year FE	Yes	Yes	No	No
Industry × Year FE	No	No	Yes	Yes
Prefecture × Year FE	No	No	Yes	Yes
Observations	63,255	63,177	63,085	63,007
R^2	0.127	0.130	0.158	0.160

Notes: The dependent variable equals 1 if the establishment hires job market entrants on net, −1 if it hires experienced workers (job market incumbents) on net, and 0 otherwise. An establishment is classified as hiring entrants (experienced workers) on net if the number of new entrants (experienced workers) hired exceeds the number of experienced workers (new entrants) hired. Standard errors in parentheses, clustered at the establishment level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

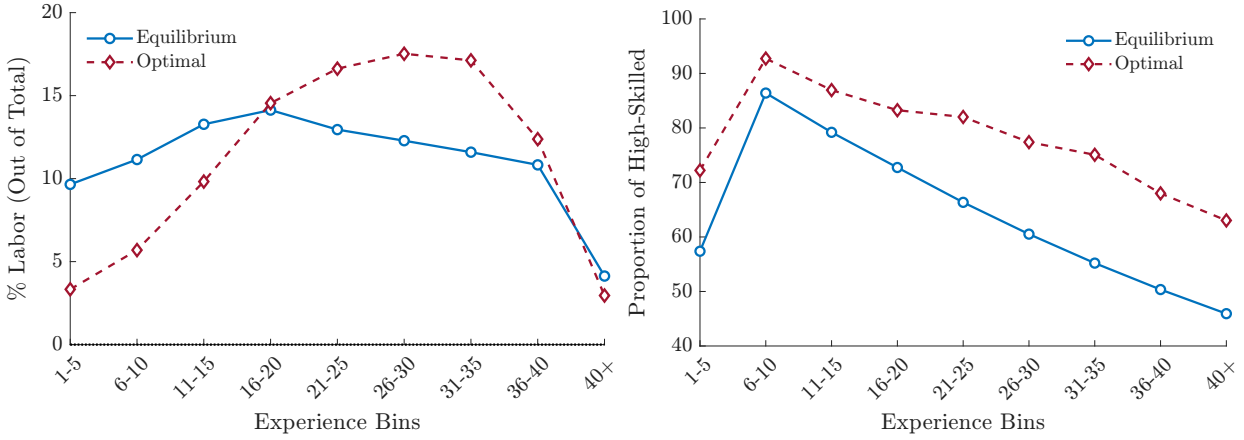
intervention.

5.1 Implementation and Policy Implications

In the baseline model with vanishing firing costs ($\tau \to 0$), the competitive equilibrium shows no within-group misallocation: labor markets clear at the group level, and firms face no adjustment frictions. The efficiency gap arises solely from the economy's inability to recompose workers across demographic groups in response to productivity differences.

This analysis provides an upper bound on the gains from workforce composition policies. In practice, interventions face additional constraints: education programs take time, retirement reforms affect only certain cohorts, fertility policies operate with decades-long lags and the fact that you must be young before you are old. Nevertheless, our quantification reveals the maximum potential gains, guiding policymakers in prioritizing interventions based on their proximity to binding productivity differences.

Figure 6: Optimal Labor Composition: Experience and Skill



(a) Optimal Experience (b) Optimal Skills

Notes: Comparison of labor composition under the competitive equilibrium (solid line with circles) versus the social planner's optimal allocation (dashed line with diamonds). Panel (a) shows the distribution across experience groups, while Panel (b) displays the fraction of high-skill workers by experience.

Figure 6 illustrates how the social planner would optimally restructure Japan's workforce to maximize aggregate output. The recomposition strategy reflects underlying productivity patterns and highlights the demographic shifts that would be most beneficial.

Panel 6a compares the experience distribution under optimal recomposition with the current equilibrium. The planner would substantially reduce the share of early-career workers (under 16 years of experience) while increasing the proportion of prime-age workers (16–40 years), consistent with the hump-shaped productivity profiles documented earlier. Workers in their peak productivity years (26–35 years) are most underrepresented in the current equilibrium, indicating significant output gains from extending careers through this period. The 40+ category is slightly overrepresented, reflecting diminishing productivity at very high experience levels.

Panel 6b examines skill composition, showing that the planner would increase the fraction of high-skill workers across most of the experience distribution. However, this increase is constrained by diminishing returns and complementarities—high-skill workers cannot be maximized indiscriminately. Notably, the gap between optimal and equilibrium skill shares widens with experience: young workers' skill composition is closer to optimal, while mid- and late-career workers face a severe deficit of high-skill labor, indicating that skill upgrading policies would be most effective for these groups.

Table 7: Output Gains from Optimal Labor Recomposition Under Different Parameters

Recomposition Scenario	**Benchmark** ($\tau \to 0$)
Optimal Experience Structure Only	1.55%
Optimal Skill Structure Only	2.38%
Optimal Joint Recomposition	3.8%

Notes: Output gains from the social planner's optimal reallocation relative to the competitive equilibrium with $\tau = 0$. "Experience Structure Only" holds the skill composition fixed at equilibrium levels while optimizing across experience groups. "Skill Structure Only" holds the experience composition fixed while optimizing the high-skill share. "Joint Recomposition" optimizes both dimensions simultaneously.

Table 7 quantifies the potential output gains from optimal labor recomposition in Japan. By jointly optimizing the experience and skill distribution of workers across firms, the social planner could raise aggregate output by 3.8%—a substantial gain achievable solely through recomposition, without any change in underlying productivities.

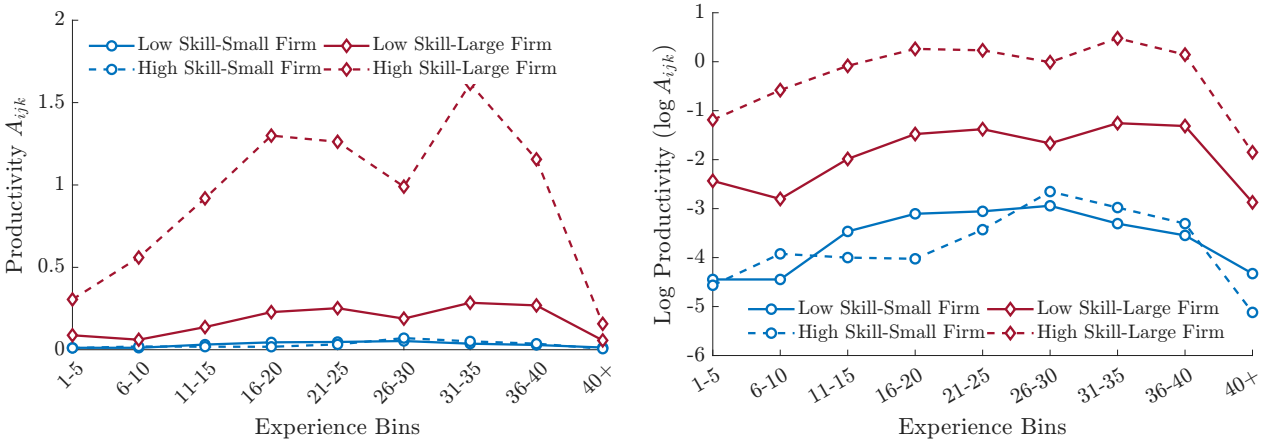
The decomposition shows that skill upgrading accounts for the majority of potential gains. Optimizing the skill composition alone would yield a 2.4% increase in output, reflecting the severe shortage of high-skill workers documented in Figure 6. The contribution of rebalancing experience or skill compositions separately is similar in magnitude when $\alpha = 0.6$, with each yielding a modest improvement in aggregate output. In contrast, rebalancing the experience distribution toward prime-age workers generates a more modest 1.6% gain. Notably, these effects are nearly additive, slightly exceeding the sum of the individual effects (3.8%), indicating a modest interaction between the two dimensions of recomposition.

These findings carry important but nuanced policy implications. While the potential gains are significant, they also reveal fundamental constraints on demographic policy. First, the dominance of skill-based gains underscores that education and retraining programs offer higher returns than policies targeting age structure alone. Second, the limited complementarity between experience and skill recomposition suggests that policymakers can pursue these objectives independently without sacrificing synergies.

5.2 Dynamic Gains to Optimal Worker Composition

Our baseline model abstracts from adjustment frictions by setting firing costs to $\tau = 0$; we examine how positive firing costs ($\tau = 0.2$) would change the estimated productivities when wages reflect both marginal products and adjustment frictions. This baseline result presents the static gain from optimal labor composition, when firms face no frictions in adjusting their labor Force. Expanding the model to include adjustment frictions implies that firms may retain workers to avoid adjustment costs. Comparing results from the static model to those from the dynamic one provides a sense of

Figure 7: Estimated Worker-Firm Productivities across Experience A_{ijk} for $\tau = 0.2$



(a) Productivity A_{ijk} (b) Log Productivity $\log A_{ijk}$

Notes: Estimated match-specific productivities for skill type $i \in \{L, H\}$, experience group $j \in \{1, ..., 9\}$, and firm type $k \in \{S, L\}$. Panel (a) shows productivity levels; panel (b) shows log productivities. Each line represents a skill-firm pair, with productivity varying across experience groups.

the costs associated with these reallocation frictions. In our baseline calibration, $\tau = 0$. We now re-estimate the model with $\tau = 0.2$ (20% of the wage payment in five years) while adjusting only the match-specific productivities A_{ijk}, keeping the remaining parameters at the benchmark levels.

Figure 7 presents the estimated productivities under positive firing costs. The key qualitative results documented above continue to hold: (i) strong complementarity between firm size and worker skill (log-supermodularity in A_{ijk}), (ii) hump-shaped productivity profiles across experience, and (iii) productivity advantages for high-skill workers that widen in large establishments. However, the productivity levels shift systematically: firing costs enlarge the estimated productivity differentials between high and low-skilled workers among experienced workers. Lower mobility of high-skilled workers due to firing costs from large to small firms results in more skilled workers in small firms than in the benchmark case. To match the wage profiles of experienced workers, the estimated productivity of high-skilled workers becomes larger in large firms.

Column 2 of Table 8 quantifies how these altered productivities affect the gains from moving to optimal composition. With $\tau = 0.2$, the total gain from optimal composition increases to 4.8% versus 3.8% in the frictionless baseline. Decomposing by margins, experience rebalancing contributes 2.4% and skill upgrading 1.9%, a more balanced split than in the baseline, where skill dominated. When there are firing costs, worker mobility over the life cycle declines. This, in turn, increases the gain from reallocating workers along the experience (i.e., age) dimension. Despite this quantitative shift, the core insight remains: Japan's workforce composition lies far from the productivity-maximizing allocation, with substantial unrealized gains available through demographic policy.

Table 8: Output Gains from Optimal Labor Composition Under Different Parameters

Parameter	**Benchmark**	$\tau = 0.2$	$\alpha = 0.4$	$\alpha = 0.8$	$\sigma = 1.5$	$\sigma = 3.5$
No Friction ($\tau = 0$)	–	0.12%	–	–	–	–
Optimal Experience Structure Only	1.55%	2.4%	0.78%	2.32%	2.41%	1.02%
Optimal Skill Structure Only	2.38%	1.9%	1.48%	2.7%	1.56%	2.78%
Optimal Joint Composition	3.8%	4.83%	2.19%	5.78%	3.54%	4.06%

Note: Output gains shown as percentage increases from current allocation. The remaining columns show robustness to alternative parameter values.

5.3 Robustness to Production Parameters

We assess the sensitivity of our main findings to key structural parameters by re-estimating worker-firm productivities A_{ijk} under alternative specifications. We first vary the returns-to-scale parameter α, and also consider alternative values of the elasticity of substitution σ.

Returns to Scale The decreasing returns parameter α governs the curvature of the firm-level production function. In this exercise, we re-estimate A_{ijk} for $\alpha \in \{0.4, 0.8\}$, bracketing our baseline $\alpha = 0.6$. Note that when we vary α, we keep σ (2.5) unchanged.

A lower α strengthens the estimated productivity advantages of workers in large firms: when returns (to production) diminish rapidly, achieving observed employment shares of large firms requires stronger underlying productivity advantages of workers in large firms. Conversely, a higher α reduces these productivity advantages. Crucially, the log-supermodularity between firm productivity and worker skill persists across all specifications, indicating this complementarity is a robust feature of the data rather than an artifact of the functional form. Appendix Figure A.7 shows the result.

Columns 3 and 4 of Table 8 substantiate how the recomposition gains vary with α. At $\alpha = 0.4$, the total gain reaches 2.2%, driven primarily by the rebalancing over the skill dimension (0.8% from experience, 1.5% from skill). At $\alpha = 0.8$, the overall gain increases to 5.8%, with a more equal split between the skill upgrading and experience rebalancing (2.3% experience, 2.7% skill). This result reveals that stronger decreasing returns (lower α) reduce the marginal benefit of reallocating any input, while weaker decreasing returns magnify these gains.[24]

Elasticity of Substitution The parameter σ controls the substitutability of different worker types in production. A smaller σ implies worker types are less substitutable, thus increasing the

[24]This finding is related the role of the returns to in scale in determining the aggregate productivity loss in the misallocation literature (e.g., Hsieh and Klenow (2009)). In that literature, the aggregate productivity loss is magnified when goods are more substitutable (i.e., weaker decreasing returns). Note that this statement is true only when the (re)estimated model parameters are largely unchanged.

value of balancing worker portfolios (conditioning on the estimated productivities). We re-estimate the model for $\sigma \in \{1.5, 3.5\}$ around the baseline $\sigma = 2.5$. We keep $\alpha = 0.6$ unchanged when changing σ.

Under a lower degree of substitutability ($\sigma = 1.5$), the estimated productivities become less dispersed across worker-firm pairs (see Appendix Figure A.8). The optimal recomposition gain reaches 3.5%, with experience rebalancing contributing disproportionately more (2.4% versus 1.6%). This reflects how lower substitutability between experience levels increases the output gains from optimally rebalancing the age distributions.

At a higher level of substitutability ($\sigma = 3.5$), the total gain from optimal labor composition reaches 4%, which is larger than the gain when $\sigma = 1.5$. The main driver is the gain from optimal skill structure. This is interesting, as gains from recomposing workers should diminish when workers become more substitutable. However, as the overall variation in the estimated productivities of various firm-worker types increases when σ goes up from 1.5 to 3.5, the gain from re-balancing the distribution of worker skills increases.

Taking Stock Across all specifications, three core findings of our quantitative exercises continue to hold: (i) worker-firm complementarity exhibits strong log-supermodularity in skill, (ii) optimal composition features substantially more prime-age (21-35 years of experience) and high-skill workers than Japan's current distribution, and (iii) potential output gains from recomposition range within 2-6%.

The decomposition between experience and skill channels shifts with parameters, but both contribute meaningfully across specifications. This suggests policies should address both margins, as neither education policy nor retirement/immigration policy alone can capture the full gains. The robustness of log-supermodularity between firm productivity and worker skill has important implications for distributional effects: population aging will systematically disadvantage high-productivity firms that rely disproportionately on younger, high-skill workers.

6 Application to Population Aging in Japan

The preceding analysis of optimal recomposition assumes the instantaneous reallocation of workers across demographic groups—a useful benchmark for quantifying potential gains, but an unrealistic depiction of feasible policy. In reality, workforce composition evolves through slow-moving demographic processes: workers accumulate experience mechanically with age, skill formation follows stochastic learning paths, and tenure patterns reflect endogenous match dynamics. These inherent dynamics impose strict feasibility constraints on compositional policies that static reallocation

exercises cannot fully capture.

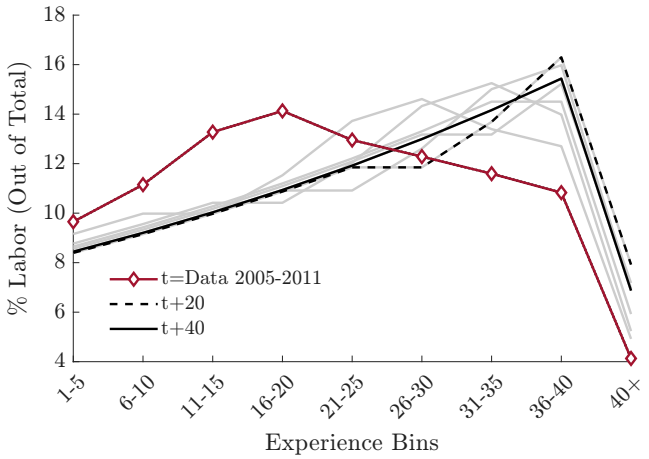
To address these limitations, we now turn to an explicitly dynamic analysis of demographic change. Specifically, we simulate the model's response to population aging over a 40-year horizon, tracking how the endogenous evolution of experience, skills, and job matches interacts with exogenous shifts in the age distribution. This exercise reveals both the direct productivity effects of demographic transition and the indirect effects operating through equilibrium adjustments in skill accumulation and worker–firm matching.

We implement this analysis by simulating cohort dynamics starting from the 2005–2011 baseline and projecting forward 40 years. To isolate composition effects from population growth, we hold the total workforce constant while allowing its age structure to evolve. The simulation incorporates Japan's below-replacement fertility rate of 1.3, implying that each successive cohort is smaller than its predecessor: $L_j/L_{j+1} = (1.3/2)^{1/5} \approx 0.92$, under the assumption that the average age of childbirth is 25 and reflecting the five-year width of our experience bins. For the oldest workers, we assume that 40% of those reaching 36–40 years of experience remain in the 40+ category, consistent with observed retirement patterns in the data. This parameterization captures Japan's rapid demographic transition while preserving tractability in the dynamic analysis. Since the labor force must remain constant over time, we proportionately adjust the sizes of the current and previous entering cohorts (starting from the initial period) every five years, while maintaining the shrinkage parameter of 8% between two consecutive entering cohorts.[25] Our interest in these exercises is the efficiency of production, where we can interpret the results as changes in output per woker in equilibrium.

We solve for the equilibrium over the 40-year horizon, assuming that firms believe the experience/skill composition of the labor supply in the current period will persist. In other words, we recompute the steady state in each period based on the contemporaneous labor supply composition. Labor and product markets clear each period, given firms' evolving beliefs about the workforce's experience and skill structure. This assumption is justified in part by the fact that a time period in the model is 5 years, so firms' longer-term expectations matter only to the extent that they base their hiring decisions on demographic trends beyond this 5-year horizon. We suspect that most firms are focused on shorter-term forecasts due to the slow pace of demographic changes and the uncertainty surrounding them. Importantly, firms' beliefs about the workforce's experience and skill structure *in future periods* do not matter in the static model where hiring/firing strategies are static choices. As a result, assumptions about firms' beliefs are irrelevant when the following

[25] Since the mass of firms is constant and does not vary with the number of workers over time, the labor supply is normalized to the mass of firms. Otherwise, under decreasing-returns technology with a fixed mass of firms, output would exhibit a level effect as the size of the labor force changes. Equivalently, we can adjust the mass of firms to be proportional to the number of workers over time. The resulting equilibrium allocations would be identical to those in our counterfactual exercises.

Figure 8: Evolution of Workforce Experience Distribution under Population Aging



Notes: Labor share by experience bin at three time points: initial (red with diamonds, 2005-2011), 20 years forward (dashed black), and 40 years forward (solid black). Based on the fertility rate of 1.3 and observed retirement patterns.

exercises are done using the static model.

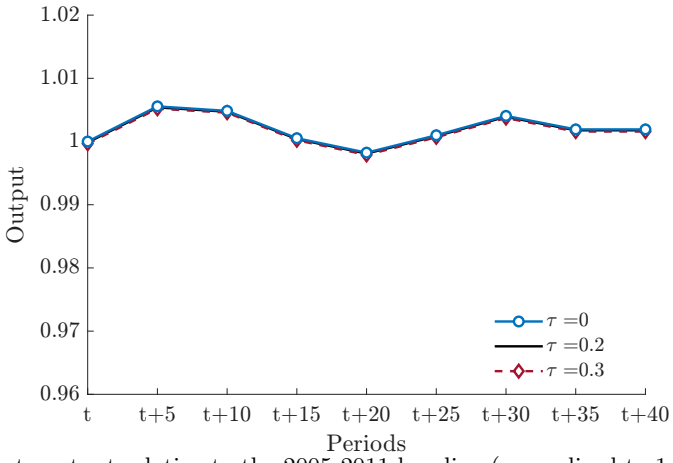
Throughout these counterfactual exercises, we use the benchmark parameters calibrated in Section 4, with firing costs set to $\tau = 0$ (equivalently, $\tau \to 0$) as in the baseline model. This frictionless benchmark allows us to isolate the demographic composition effects on aggregate outcomes. We also explore the sensitivity of our results to positive firing costs by conducting robustness checks with $\tau > 0$, holding all other parameters fixed at their baseline calibrated values. These exercises reveal how labor market frictions interact with demographic change to shape the economy's response to population aging.

Figure 8 shows the projected evolution of Japan's workforce composition under our demographic assumptions. Starting from the 2005–2011 baseline (red diamonds), we simulate how the experience distribution evolves when each cohort is 8% smaller than its predecessor. The 20-year projection (dashed line) reveals a flattening of the distribution, as smaller entry cohorts partially offset retirements. By year 40 (solid line), the full demographic shift materializes: the 36–40 category rises to nearly 15% of workers, while those with less than 15 years of experience fall below 10%. These projections form the basis for our counterfactual analysis of how aging affects aggregate output, wages, and firm concentration through both compositional and equilibrium channels.

6.1 Population Aging, Productivity, and Firm Concentration

We examine two dimensions of population aging's impact: its effects on aggregate productivity and its consequences for the distribution of workers across firms. The analysis reveals distinct patterns

Figure 9: Effect of Population Aging on Aggregate Output



Notes: Evolution of aggregate output relative to the 2005-2011 baseline (normalized to 1.0) under population aging. Lines show different firing cost scenarios: $\tau = 0$ (no firing costs), $\tau = 0.2$ (20% of wages), and $\tau = 0.3$ (30% of wages). The x-axis represents years forward from the initial period.

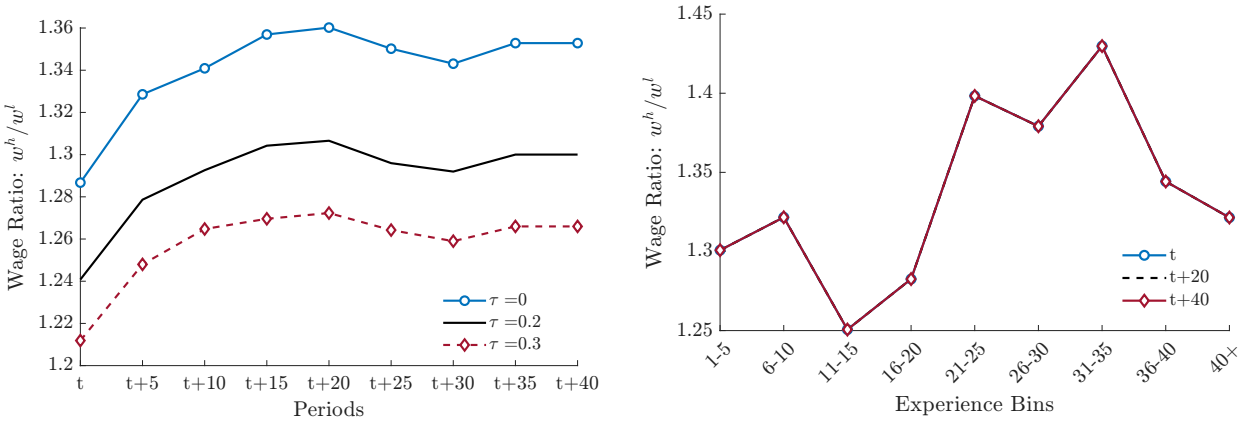
at both the aggregate and distributional levels.

Figure 9 presents the evolution of aggregate output under a demographic transition. Population aging has only a modest impact over the 40-year projection horizon, with the maximum negative and positive effects around −1% and 1%, respectively. This reflects the fact that the age–skill composition of the labor force remains far from the optimal allocation throughout the period. Specifically, there are too few workers in the four experience bins (16–35 years) compared with the optimal allocation in Figure 6, where each bin is at or above 15%. Introducing firing costs of up to 30% of wages ($\tau = 0.3$) has only a minimal effect on this trajectory, indicating that labor market frictions play a limited role in mediating demographic effects on aggregate productivity.

Turning to the distributional effects, Figure 10 illustrates how population aging influences wage inequality across skill groups. Panel (a) traces the skill premium—the percentage wage advantage of high-skill workers over low-skill workers—as it rises from 29% to 36% over the projection period when $\tau = 0$. This 7-percentage-point increase corresponds to a 24.1% rise in relative wage inequality. This increase is driven primarily by compositional shifts—specifically, the growing share of more experienced workers, who command a higher skill premium—rather than by equilibrium wage adjustments.

Panel (b) decomposes this pattern by experience groups, highlighting the underlying compositional mechanism. We focus on the frictionless benchmark case with no firing cost ($\tau = 0$). The skill premium increases systematically with experience, from roughly 30% for early-career workers to more than 40% for senior cohorts. Within each experience bin, however, the wage gap remains essentially

Figure 10: Skill Premium Evolution under Population Aging



(a) Skill Premium Over Time

(b) Skill Premium by Experience ($\tau = 0$)

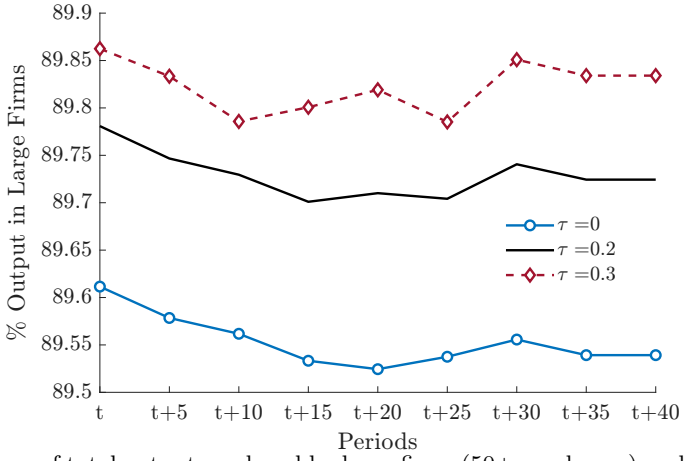
Notes: Panel (a) shows the evolution of the aggregate skill premium (high-skill/low-skill wage ratio) under different firing cost scenarios. Panel (b) displays skill premiums by experience group at three time points: initial (solid), 20 years forward (dashed), and 40 years forward (dotted). Experience bins are in 5-year intervals.

constant over time (as indicated by the nearly overlapping lines), suggesting only limited equilibrium adjustments in relative wages. The increase in the aggregate skill premium, therefore, reflects changes in the workforce composition (towards more senior workers with higher skill premium) rather than shifts in the returns to skill.

Firing costs primarily affect the level of the skill premium. When $\tau = 0.3$, the premium is about 8 percentage points lower than in the frictionless benchmark. As τ rises, it becomes increasingly difficult for large (small) firms to dismiss unskilled (skilled) workers. As a result, skilled workers become trapped in low-productivity small firms, while unskilled workers who would have been dismissed in a frictionless environment remain employed in high-productivity large firms. This mismatch between workers' skills and firms' productivity compresses the skill premium. Notably, the change in inequality induced by population aging is similar across friction levels, indicating that firing costs shift the level (but not the trajectory) of wage dispersion.

Figure 11 examines how population aging and firing costs affect the firm-size distribution by tracking the share of aggregate output produced by large firms (those with 50 or more employees). First, the concentration ratio of top firms generally declines as the population ages. The intuition is that productive firms rely disproportionately on young and skilled workers. As the labor force becomes older—and therefore, on average, less skilled—the relative advantage of productive firms diminishes, shifting output toward less productive firms and reducing the market share of large, high-productivity producers. This mechanism yields a key testable prediction of the model. We present empirical evidence consistent with this prediction in Section 6.2.

Figure 11: Large Firm Output Share under Population Aging



Notes: Evolution of the share of total output produced by large firms (50+ employees) under population aging. Lines show different firing cost scenarios.

The figure also shows that higher firing costs increase the concentration ratio of top firms. This effect emerges through equilibrium adjustments in the skill premium. When firing costs rise, the skill premium declines. Because productive firms employ relatively more skilled workers, they benefit from this reduction in skilled labor costs. As a result, their profitability improves, and their market share expands.

6.2 Evidence: Population Growth/Aging and Market Concentration

In this subsection, we test our model's prediction that population aging reduces market concentration. Because more productive firms rely relatively more on worker skill than experience, a shift in labor supply toward older workers disproportionately harms them, lowering their market share. To examine this, we exploit variation in the growth rate of the working-age population across 47 Japanese prefectures and study its effect on the market shares of the top 5% and 10% of establishments at both the prefecture–year and prefecture–industry–year levels.[26] For the prefecture–year (prefecture–industry–year) analysis, we compute the sales and employment shares of the top 5% and 10% of establishments within each unit and year. The prefecture–industry–year specification helps control for compositional changes,[27] but some cells contain too few establishments to compute market shares. We therefore report results at both levels and restrict the prefecture–industry–year

[26]We focus on within-country, across-region variation for two reasons. First, there is substantial heterogeneity in working-age population growth across Japanese prefectures over our sample period. The average prefectural growth rate fell from 0.7% in 1985 to −1.3% in 2018, while the standard deviation declined only slightly, from 0.8% to 0.6%. Second, this design mitigates concerns that arise in cross-country studies, as regions within the same country (particularly Japan) are subject to broadly similar policies, institutions, and culture.

[27]E.g., prefectures shifting toward industries that are intrinsically more concentrated.

sample to cells with at least 100 establishments.[28]

We use a confidential establishment-level dataset together with official statistics on prefecture-level demographic variables from Japan to conduct the analysis. The establishment-level dataset, the Census of Manufacturers (1985–2018), is mandatory for all manufacturing establishments and is provided by the Japanese Ministry of Economy, Trade, and Industry. It surveys all establishments with more than three employees annually, generating a panel spanning over 30 years and covering approximately 300,000–350,000 establishments per year.[29] From this dataset, we compute the sales and employment shares of the top 5% and 10% of establishments.

For prefecture-level demographic variables, we obtain data on the growth rate of the working-age population and the average age of the population from the National Population Census (1985–2018), conducted by the Japanese Ministry of Internal Affairs and Communications. We then merge the two datasets and estimate the following panel regression:

$$concentration_{p,y}^{top\ firm} = \beta_0 + \beta_1 working\ gr_{p,y} + \delta_p + \delta_y + \epsilon_{p,y}, \tag{14}$$

where $concentration_{p,y}^{top\ firm}$ is the sales/employment shares of top 5% (or 10%) establishments in prefecture p and year y at either the prefecture-year level or the prefecture-industry-year level. The variable of $working\ gr_{p,y}$ is the annual growth rate of the working-age population measured at the prefecture-year or prefecture-industry-year level. We always include the prefecture fixed effects δ_p (industry-prefecture fixed effects δ_{ip}) and the year fixed effects into the prefecture-year (prefecture-industry-year) regression. As a result, we are comparing prefectures/industry-prefecture pairs with different growth rates of working age population in the same time period. Standard errors are clustered at the year and prefecture levels for the prefecture-year level regressions (i.e., two-way clustering) and at the prefecture-year level for the prefecture-industry-year level regressions.

Table 9 supports our model's prediction: β_1 is positive and significant across all specifications, indicating that the employment share of top establishments drops in prefectures with slower growth of working-age population. This result is robust to different top-firm thresholds and concentration measures (see Table B.7). Using average population age as a proxy for working-age population age yields similar results: prefectures with faster aging populations have lower market shares for the largest establishments (Table B.8). Overall, these findings provide empirical support that slower population growth or rapid aging reduces market concentration and the dominance of top firms.

[28] Results are robust to a 50-establishment threshold.

[29] In years ending in 0, 2, 5, and 8, the Census of Manufacturers surveys all manufacturing establishments.

Table 9: Working-age population growth and employment concentration

	Prefecture-Year		Industry-Prefecture-Year	
	(1) Top 5%	(2) Top 10%	(3) Top 5%	(4) Top 10%
Working-age pop. growth	1.458***	1.211***	0.464***	0.391***
	(0.427)	(0.372)	(0.108)	(0.102)
Prefecture FE	Yes	Yes	No	No
Industry × Prefecture FE	No	No	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	1,598	1,598	19,913	19,913
Adjusted R^2	0.894	0.902	0.730	0.733

Notes: The dependent variable is the employment share of the largest establishments (top 5% or top 10% by employment) in each prefecture-year (columns 1–2) or industry-prefecture-year (columns 3–4). Working-age population growth is measured at the prefecture-year level. Standard errors are two-way clustered by prefecture and year (columns 1–2) or clustered at the prefecture-year level (columns 3–4). Columns (3)–(4) include only industry-prefecture-year cells with at least 100 establishments. Data from Census of Manufacturing Plants, 1985–2018. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Confounding Factors: Establishment Age and Worker Migration

Establishment and firm age is positively correlated with market share, as surviving establishments/-firms tend to grow. Thus, average age is a potential confounder in our regressions. Population aging or slower working-age population growth leads to older firms, which would suggest higher market concentration—opposite to our empirical finding. Thus, this confounding factor works *against* our results. Controlling for it would likely make the positive effect of population growth (and negative effect of aging) on market concentration even stronger.

Worker mobility may induce endogeneity, as shocks that expand large firms could also attract young workers, creating reverse causality. To address this, we follow Maestas et al. (2016) and Karahan et al. (2024) and use the 20-year lagged birth rate at the prefecture-year level as an instrumental variable for the annual growth rate of the working-age population. Since many young adults become part of the working-age population in their prefecture of birth, the prefecture's lagged birth rate will influence the future growth of its working-age population. The exclusion restriction is that, conditional on prefecture and time fixed effects, the determinants of a higher birth rate have no effect on future establishments except indirectly through their impact on the growth of the working-age population.[30]

[30]This assumption would be violated, for instance, if people in a given prefecture had a higher fertility rate 20 years ago in anticipation of stronger labor market conditions (relative to other prefectures). However, as local economies in Japan (and the national economy) experienced significant fluctuations between 1975 and 2018, and given the long 20-year time lag, it is unlikely that individual households had such perfect foresight when making their fertility

Table 10: Population growth and employment concentration: IV estimates using lagged birth rates

	First Stage	Prefecture-Year		Industry-Prefecture-Year	
	(1)	(2)	(3)	(4)	(5)
	Working-Age Pop. Growth	Employment Share Top 5%	Top 10%	Employment Share Top 5%	Top 10%
Working-age pop. growth		5.007***	4.325***	1.607***	1.645***
		(1.754)	(1.492)	(0.258)	(0.239)
Birth rate$_{t-20}$	0.124***				
	(0.042)				
F-statistic (first stage)	284.1				
Prefecture FE	Yes	Yes	Yes	No	No
Industry × Prefecture FE	No	No	No	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Observations	1,588	1,588	1,588	19,873	19,873
R^2	0.908	0.866	0.876	0.744	0.747
Adjusted R^2	0.903	0.859	0.869	0.729	0.733

Notes: Column (1) presents the first-stage regression of working-age population growth on the 20-year lagged birth rate at the prefecture level. Columns (2)–(5) present second-stage IV estimates where the dependent variable is the employment share of the largest establishments (top 5% or top 10% by employment) in each prefecture-year (columns 2–3) or industry-prefecture-year (columns 4–5). Working-age population growth is instrumented using the 20-year lagged prefecture birth rate. Standard errors in parentheses are two-way clustered by prefecture and year (columns 1–3) or clustered at the prefecture-year level (columns 4–5). Columns (4)–(5) include only industry-prefecture-year cells with at least 100 establishments with non-missing sales and employment. Data from Census of Manufacturing Plants, 1985–2018. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 10 confirms the results reported in Table 9.[31] The effect is quantitatively sizable: a 0.86% decline in working-age population growth (one standard deviation) reduces the employment share of the top 10% of establishments by 3.68% at the prefecture-year level and 1.40% at the industry-prefecture-year level (standard deviations: 5.21% and 13.0%, respectively). In summary, both OLS and IV results support our model's key prediction: slower growth of the working-age population and higher average worker age reduce the employment share of top establishments.

decisions two decades ago.

[31] The correlation between the working-age population growth rate and the 20-year lagged birth rate is 0.87 in our sample.

7 Conclusion

As populations age across developed economies, understanding how demographic shifts reshape labor force composition and its consequences for productivity, inequality, and firm dynamics has become increasingly important. This paper develops a framework to characterize the optimal allocation of workers across heterogeneous firms and quantifies the economic effects of population aging using Japanese establishment-worker data.

We document that larger, more productive establishments employ younger workers with longer tenure, revealing complementarities between firm productivity and worker skill that generate differential exposure to demographic change. Our main quantitative finding is that reallocating workers toward the optimal composition would increase Japan's GDP by 3.8%. Incorporating dynamics and labor market frictions, we show that population aging produces heterogeneous effects across the economy: it increases wage inequality between skill groups while reducing market concentration as high-productivity firms lose their comparative advantage. Prefecture-level evidence validates these predictions.

Our findings carry important policy implications. The dominance of skill-based gains suggests that education and retraining programs targeting mid-career workers offer particularly high returns in aging societies. Because experience and skill recomposition operate relatively independently, policymakers can pursue these objectives separately. Employment protection policies generate distributional effects that vary systematically across firm productivity levels and worker types, highlighting the need for labor market institutions that facilitate efficient reallocation while accounting for firm heterogeneity and worker dynamics.

An exciting direction for future research is to endogenize technological responses to demographic change. Our analysis holds production technology fixed, but aging populations may induce firms to adopt automation that substitutes for scarce young workers or innovations that better leverage older worker experience. Such endogenous technological adaptation could either reinforce or mitigate the productivity gaps we identify. Understanding these responses would provide a more complete picture of how economies adjust to demographic transitions and reveal whether technological change ultimately amplifies or dampens the effects of population aging on aggregate outcomes.

References

Alvarez, Fernando and Marcelo Veracierto, "Severance payments in an economy with frictions," *Journal of monetary Economics*, 2001, *47* (3), 477–498.

Bagga, Sadhika, "Internal and External Labor Markets and Declining Dynamism," 2022.

Bentolila, Samuel and Giuseppe Bertola, "Firing costs and labour demand: how bad is eurosclerosis?," *The Review of Economic Studies*, 1990, *57* (3), 381–402.

Berger, Mark C, "The Effect of Cohort Size on Earnings Growth: A Reexamination of the Evidence," *Journal of Political Economy*, June 1985, *93* (3), 561–573.

Bianchi, Nicola and Matteo Paradisi, "Countries for Old Men: An Analysis of the Age Pay Gap," Technical Report, National Bureau of Economic Research 2024.

Borjas, George J, "The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market," *The quarterly journal of economics*, 2003, *118* (4), 1335–1374.

Braun, R. Anton and Douglas H. Joines, "The implications of a graying Japan for government policy," *Journal of Economic Dynamics and Control*, None 2015, *57* (C), 1–23.

_ , Daisuke Ikeda, and Douglas H. Joines, "The Saving Rate In Japan: Why It Has Fallen And Why It Will Remain Low," *International Economic Review*, February 2009, *50* (1), 291–321.

Card, David and Thomas Lemieux, "Can falling supply explain the rising return to college for younger men? A cohort-based analysis," *The quarterly journal of economics*, 2001, *116* (2), 705–746.

Cardona, Carolina, "How a Vast Demographic Shift Will Reshape the World," *The New York Times*, 2023.

Engbom, Niklas et al., "Firm and worker dynamics in an aging labor market," Technical Report 2019.

Frey, William H., "Immigration drives the nation's healthy post-pandemic population growth, new census data show," Technical Report, Brookings Insitution 2025.

Geruso, Michael and Dean Spears, "The Likelihood of Persistently Low Global Fertility," *Journal of Economic Perspectives*, Winter 2026, *40* (1), 3–26.

Gregory, Victoria, "Firms as learning environments: Implications for earnings dynamics and job search," *FRB St. Louis Working Paper*, 2020, (2020-036).

Hayashi, Fumio, "The Root Cause of Japan's 30-year Stagnation: Implications for the US, Europe, and East Asia," *Europe, and East Asia (August 28, 2023)*, 2023.

Hopenhayn, Hugo A, "Entry, exit, and firm dynamics in long run equilibrium," *Econometrica: Journal of the Econometric Society*, 1992, pp. 1127–1150.

Hopenhayn, Hugo and Richard Rogerson, "Job turnover and policy evaluation: A general equilibrium analysis," *Journal of political Economy*, 1993, *101* (5), 915–938.

_ , Julian Neira, and Rish Singhania, "From population growth to firm demographics: Im-

plications for concentration, entrepreneurship and the labor share," *Econometrica*, 2022, *90* (4), 1879–1914.

Hsieh, Chang-Tai and Peter J Klenow, "Misallocation and manufacturing TFP in China and India," *The Quarterly journal of economics*, 2009, *124* (4), 1403–1448.

IMF, "Japan: Demographic Shift Opens Door to Reforms," Technical Report, IMF Country Focus 2020.

Karahan, Fatih, Benjamin Pugsley, and Ayşegül Şahin, "Demographic Origins of the start-up Deficit," *American Economic Review*, 2024, *114* (7), 1986–2023.

Katz, Lawrence F and Kevin M Murphy, "Changes in relative wages, 1963–1987: supply and demand factors," *The quarterly journal of economics*, 1992, *107* (1), 35–78.

Kim, Seula, "Workers' Job Prospects and Young Firm Dynamics," Technical Report, Working Paper 2023.

Kitao, Sagiri, "Fiscal cost of demographic transition in Japan," *Journal of Economic Dynamics and Control*, None 2015, *54* (C), 37–58.

Maestas, Nicole, Kathleen J Mullen, and David Powell, "The effect of population aging on economic growth, the labor force and productivity," Technical Report, National Bureau of Economic Research 2016.

Melitz, Marc J, "The impact of trade on intra-industry reallocations and aggregate industry productivity," *econometrica*, 2003, *71* (6), 1695–1725.

Ottaviano, Gianmarco IP and Giovanni Peri, "Rethinking the effect of immigration on wages," *Journal of the European economic association*, 2012, *10* (1), 152–197.

Ouimet, Paige and Rebecca Zarutskie, "Who works for startups? The relation between firm age, employee age, and growth," *Journal of financial Economics*, 2014, *112* (3), 386–407.

Papageorgiou, Theodore, "Large firms and within firm occupational reallocation," *Journal of Economic Theory*, None 2018, *174* (C), 184–223.

Peters, Michael and Conor Walsh, "Population growth and firm dynamics," Technical Report, National Bureau of Economic Research 2021.

Poschke, Markus, "Employment protection, firm selection, and growth," *Journal of Monetary Economics*, 2009, *56* (8), 1074–1085.

Qi, Liyan, "China Sees a Fresh Decline in Population, Despite a Rise in Births," *The Wall Street Journal*, 2025.

Toossi, Mitra and Torpey Elka, "Older workers: Labor force trends and career options," Technical Report, Bureau of Labor Statistics 2017.

Veracierto, Marcelo, "Employment flows, capital mobility, and policy analysis," *International Economic Review*, 2001, *42* (3), 571–596.

Weil, David N., "How Much Would Continued Low Fertility Affect the US Standard of Living?," *Journal of Economic Perspectives*, Winter 2026, *40* (1), 27–46.

Welch, Finis, "Effects of Cohort Size on Earnings: The Baby Boom Babies' Financial Bust," *Journal of Political Economy*, 1979, *87* (5), S65–97.

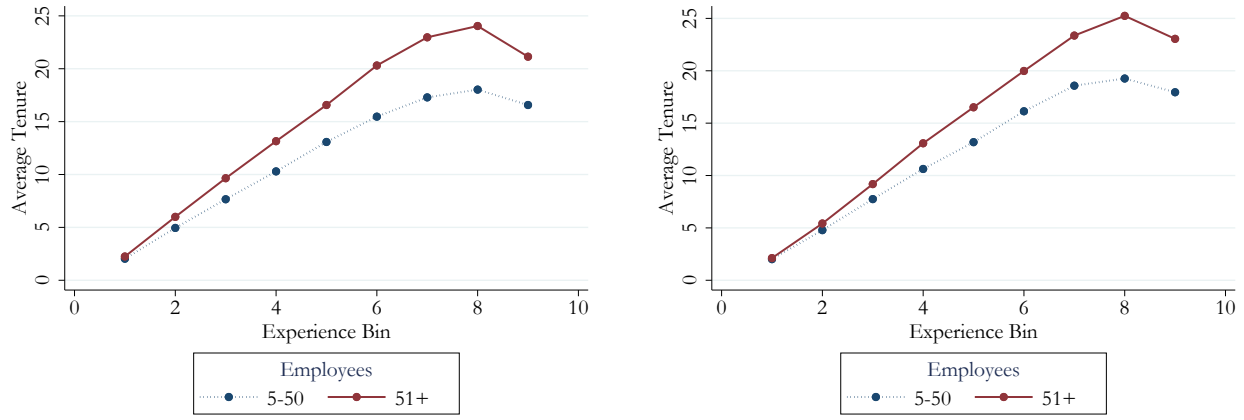
Appendix

This appendix provides supplementary empirical evidence, theoretical foundations, and computational details for the main analysis. Section A presents additional empirical patterns from Japanese labor market data. Section B provides detailed summary statistics and regression results. Section C contains formal proofs of theoretical propositions, efficiency analysis, and identification arguments.

A Additional Empirical Evidence: Figures

This section presents additional empirical patterns that complement the main text, including robustness checks with different time periods, sample restrictions, and alternative firm size classifications.

Figure A.1: Experience-Tenure Profiles by Period and Worker Type

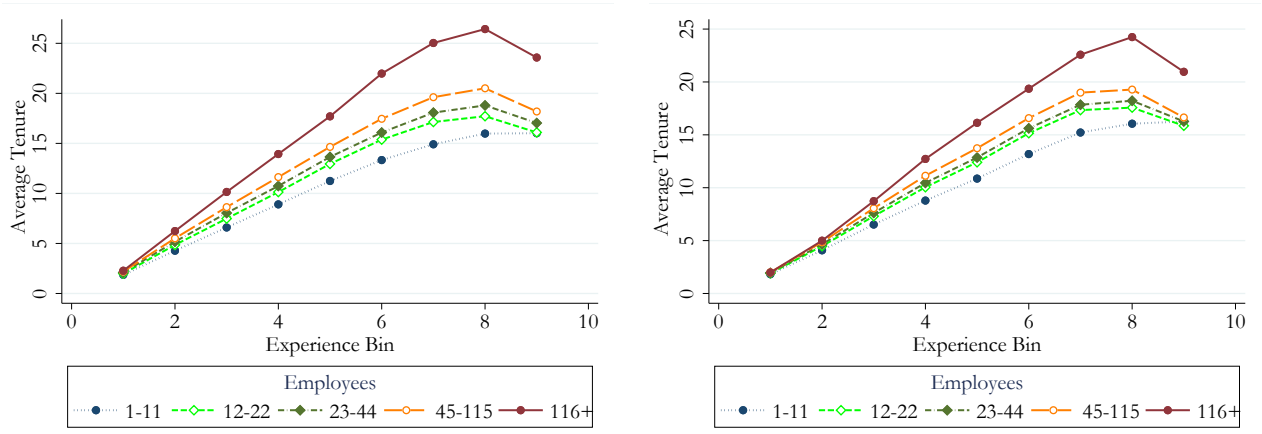


(a) 1998–2004

(b) 2005–2011: Permanent Workers

Notes: Experience is defined as age minus (length of education + 6). Experience bins are defined as: bin 1 (1–5 years), bin 2 (6–10 years), ..., bin 8 (36–40 years), bin 9 (41–45 years). The experience bins are capped at 45 years. Small firms have fewer than 50 employees; large firms have 50 or more employees.

Figure A.2: Experience-Tenure Profiles: Five Size Bins

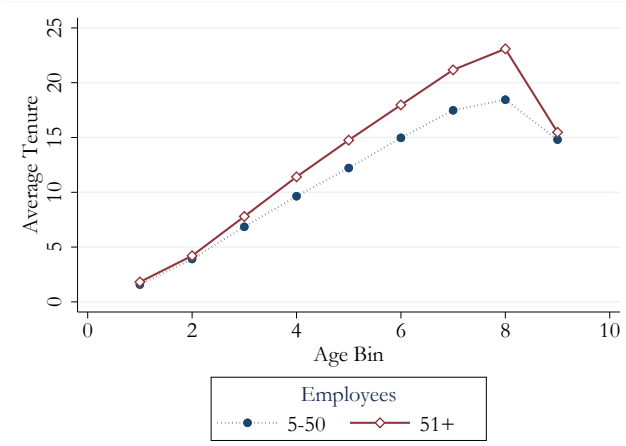


(a) 1998–2004

(b) 2005–2011

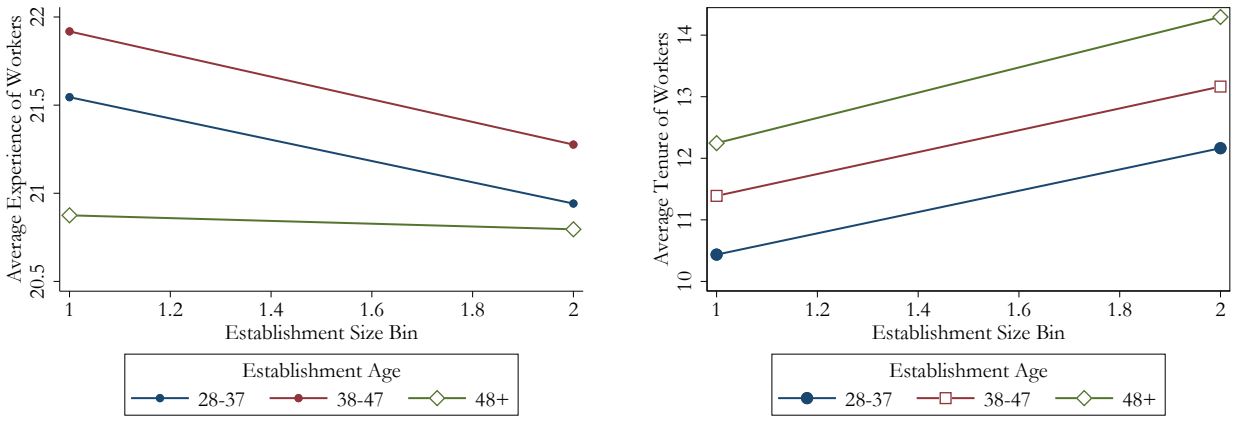
Notes: Experience is defined as age minus (length of education + 6). Experience bins are defined as: bin 1 (1–5 years), bin 2 (6–10 years), ..., bin 8 (36–40 years), bin 9 (41–45 years). The experience bins are capped at 45 years. Firm size bins are defined as: bin 1 (≤11 employees), bin 2 (12–22 employees), bin 3 (23–44 employees), bin 4 (45–115 employees), bin 5 (≥116 employees).

Figure A.3: Age-tenure profile (2005-2011)



Note: Age bin 1: 20-24 years old; Age bin 2: 25-29 years old; ... Age bin 9: 60-64 years old. Small establishments have fewer than 50 employees; large establishments have 50 or more employees.

Figure A.4: Average Experience and Tenure by Establishment Age and Size (2005–2011)

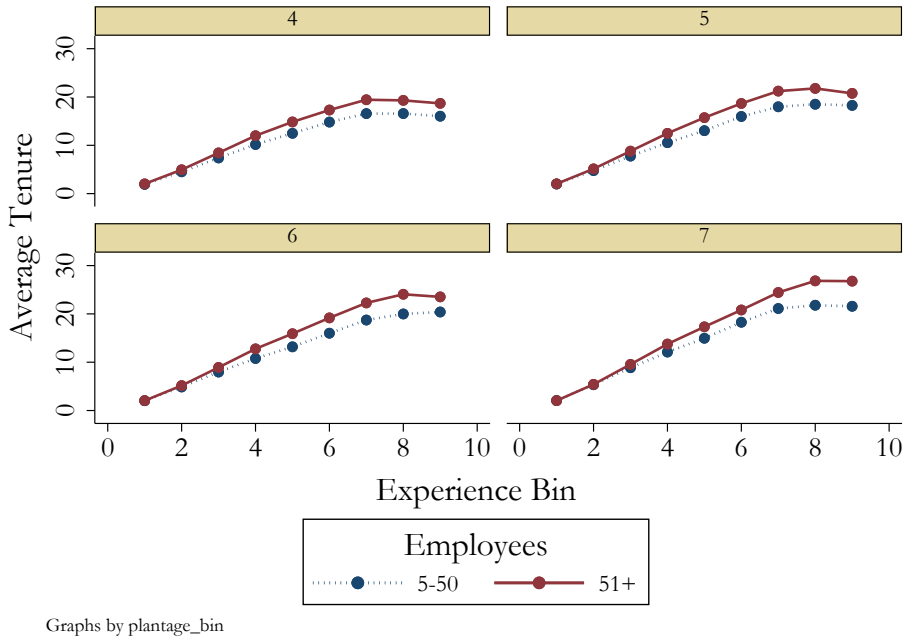


(a) Average Experience

(b) Average Tenure

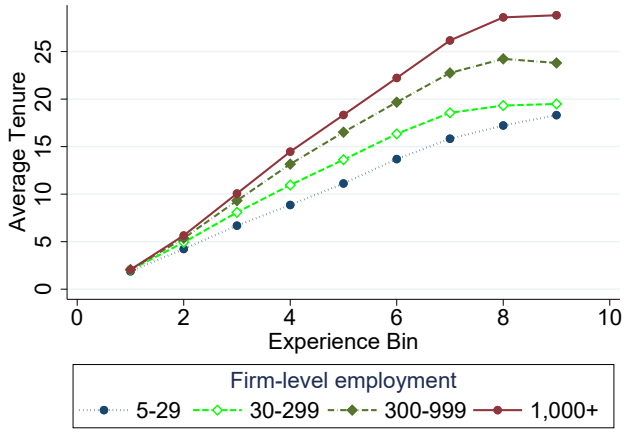
Notes: Average experience (panel a) and tenure (panel b) by establishment age and size. Sample restricted to workers before retirement age. Small establishments have fewer than 50 employees; large establishments have 50 or more employees.

Figure A.5: Experience-tenure profile by establishment age (2005-2011): workers before retirement



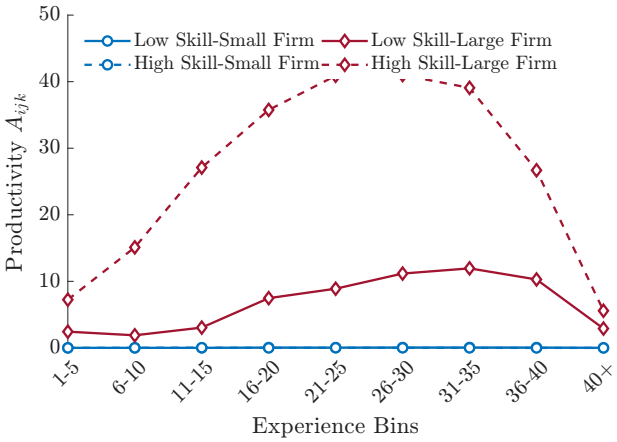
Note: Experience = age - (length of education + 6). Experience bin 1: 1-5 years; ... Experience bin 9: 41-45 years. We cut the experience bin at 45 years. Age bin 4: establishment age 18-27; Age bin 5: establishment age 28-37; Age bin 6: establishment age 38-47; Age bin 7: establishment age 48+

Figure A.6: Experience-Tenure Profile at the Firm Level: 2005-2011

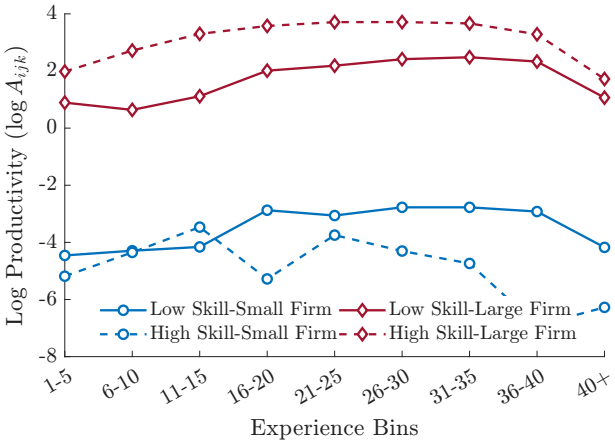


Note: Establishments founded after 1985 are excluded, and workers older than 60 are excluded. Experience = age - (length of education + 6). Experience bin 1: 1-5 years; ... Experience bin 8: 36-40 years; Experience bin 9: 41-45 years.

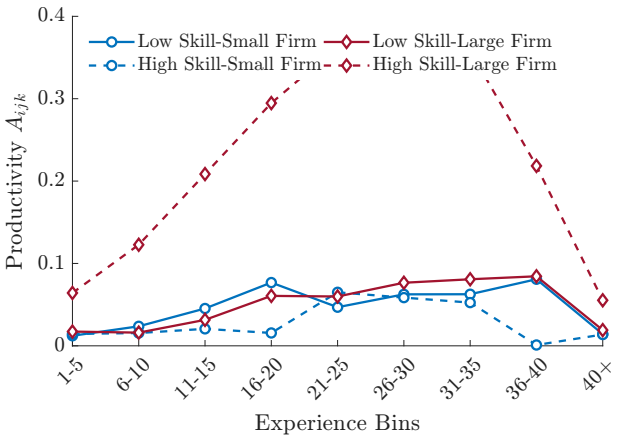
Figure A.7: Estimated Worker-Firm Productivities across Experience A_{ijk} Varying α



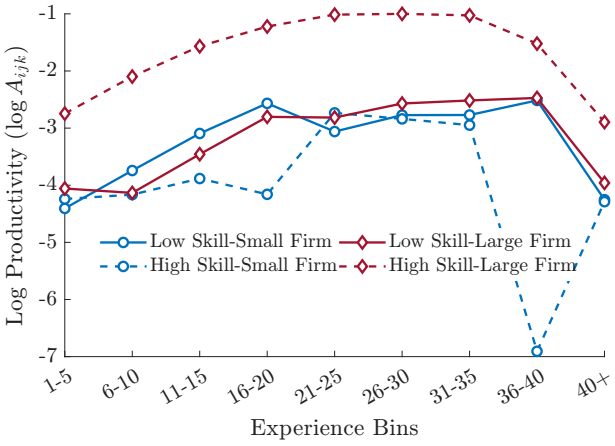
(a) Productivity A_{ijk}, $\alpha = 0.4$



(b) Log Productivity $\log A_{ijk}$, $\alpha = 0.4$



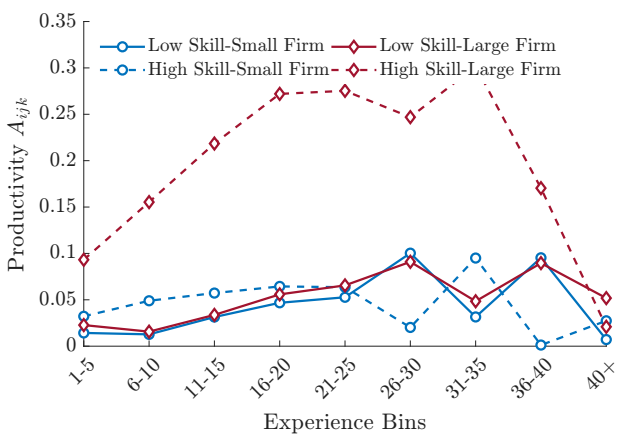
(c) Productivity A_{ijk}, $\alpha = 0.8$



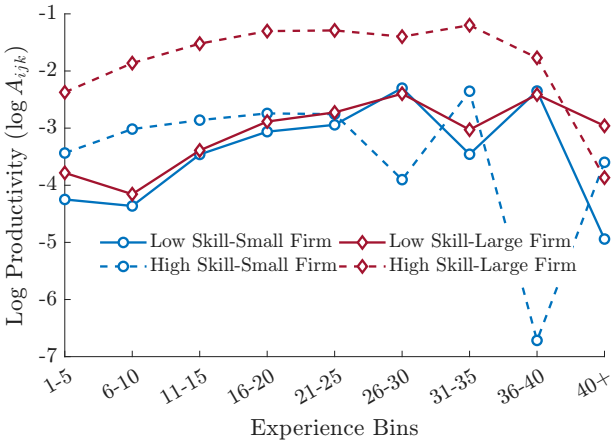
(d) Log Productivity $\log A_{ijk}$, $\alpha = 0.8$

Notes: Estimated match-specific productivities for skill type $i \in \{L, H\}$, experience group $j \in \{1, ..., 9\}$, and firm type $k \in \{S, L\}$. Panel (a) shows productivity levels; panel (b) shows log productivities. Each line represents a skill-firm pair, with productivity varying across experience groups.

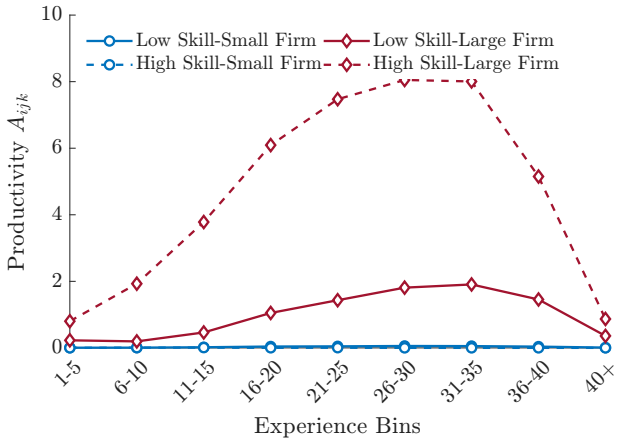
Figure A.8: Estimated Worker-Firm Productivities across Experience A_{ijk} Varying σ



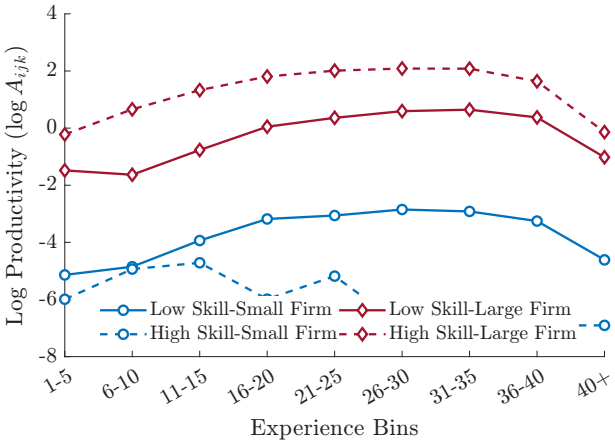
(a) Productivity A_{ijk}, $\sigma = 1.5$



(b) Log Productivity $\log A_{ijk}$, $\sigma = 1.5$



(c) Productivity A_{ijk}, $\sigma = 3.5$



(d) Log Productivity $\log A_{ijk}$, $\sigma = 3.5$

Notes: Estimated match-specific productivities for skill type $i \in \{L, H\}$, experience group $j \in \{1, ..., 9\}$, and firm type $k \in \{S, L\}$. Panel (a) shows productivity levels; panel (b) shows log productivities. Each line represents a skill-firm pair, with productivity varying across experience groups.

B Additional Empirical Evidence: Tables

This section provides detailed summary statistics and regression results that support the empirical findings in the main text, including cross-sectional regressions of worker composition by establishment characteristics.

Table B.1: Average Worker Characteristics by Two Size Bins (1998-2004): full sample

Size Bin	Employment	Age	Experience	Tenure	Wage	Education (years)
1	24.71	40.69	22.01	11.20	35748.69	12.68
2	442.15	39.47	20.30	13.70	42271.18	13.17

Note: Time period: 1998-2004. Unit of average wage: one hundred yen (or one USD). We divide establishments to two sized bins. Size bin 1: employment ≤ 50; Size bin 2: employment > 50.

Table B.2: Average Worker Characteristics by Two Size Bins (2005-2011): permanent workers only

Size Bin	Employment	Age	Experience	Tenure	Wage	Education
1	23.62	40.76	21.50	11.60	36422.90	13.25
2	418.83	40.18	20.56	14.03	43197.00	13.62

Note: Time period: 2005-2011. Unit of average wage: one hundred yen (or one USD). We divide establishments to two sized bins. Size bin 1: employment ≤ 50; Size bin 2: employment > 50.

Table B.3: Average Worker Characteristics by Five Size Bins (2005-2011): full sample

Size Bin	Employment	Age	Experience	Tenure	Wage	Education (years)
1	8.53	42.58	23.81	10.57	31221.04	12.77
2	16.75	40.77	21.56	10.68	34566.54	13.21
3	31.81	40.92	21.68	11.10	35674.05	13.24
4	73.14	41.10	21.79	11.81	37151.72	13.31
5	620.85	40.41	20.83	13.46	42481.81	13.58

Note: Time period: 2005-2011. Unit of average wage: one hundred yen (or one USD). We divide establishments to five size equal-sized bins. Size bin 1: employment ≤ 11; Size bin 2: employment $\in [12, 22]$; Size bin 3: employment $\in [23, 44]$; Size bin 4: employment $\in [45, 115]$; Size bin 5: employment ≥ 116.

Table B.4: Average Worker Characteristics by Five Size Bins (1998-2004): full sample

Size Bin	Employment	Age	Experience	Tenure	Wage	Education (years)
1	8.48	42.15	23.81	10.55	32056.02	12.34
2	16.84	40.61	21.92	10.95	35683.63	12.69
3	32.17	40.32	21.55	11.51	36838.49	12.77
4	72.45	40.06	21.17	12.28	38261.13	12.89
5	665.22	39.12	19.80	14.51	44632.34	13.32

Note: Time period: 1998-2004. Unit of average wage: one hundred yen (or one USD). We divide establishments to five size equal-sized bins. Size bin 1: employment ≤ 11; Size bin 2: employment $\in [12, 22]$; Size bin 3: employment $\in [23, 44]$; Size bin 4: employment $\in [45, 115]$; Size bin 5: employment ≥ 116.

Table B.5: Cross-sectional regression of experience, age and tenure: 2004-2011

Dep.Var:	(1) ave. experience	(2) ave. age	(3) ave. tenure
log(employment)	-0.355***	-0.310***	0.531***
	(0.011)	(0.011)	(0.011)
establishment age	0.056***	0.056***	0.082***
	(0.001)	(0.001)	(0.001)
constant	49.275***	53.854***	8.152***
	(0.167)	(0.151)	(0.119)
city FE	Yes	Yes	Yes
Industry-year FE	Yes	Yes	Yes
Prefecture-year FE	Yes	Yes	Yes
N	365721	365721	365721
R^2	0.400	0.398	0.454

* 0.10 ** 0.05 *** 0.01. Standard errors are clustered at the establishment level.

Table B.6: Worker Characteristics and Firm Size: 2005-2011

Size Bin	Experience	Tenure	Wage
1	22.94	10.04	27686.5
2	21.61	11.28	31746.4
3	20.22	12.81	36279.1
4	19.91	13.98	42235.4

Time period: 2005-2011. *exp*: experience (years of working). Unit of *wage*: one hundred yen (or one USD). Establishments founded after 1985 are excluded, and workers older than 60 are excluded. We divide firms to four size bins. Size bin 1: emp $\in$ [5,29]; Size bin 2: emp $\in [30, 299]$; Size bin 3: emp $\in [300, 999]$; Size bin 4: emp ≥ 1000.

Table B.7: Working-age population growth and sales concentration

	Prefecture-Year		Industry-Prefecture-Year	
	(1) Top 5%	(2) Top 10%	(3) Top 5%	(4) Top 10%
Working-age pop. growth	1.722***	1.229***	0.309**	0.250**
	(0.522)	(0.381)	(0.145)	(0.122)
Prefecture FE	Yes	Yes	No	No
Industry × Prefecture FE	No	No	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	1,598	1,598	19,913	19,913
Adjusted R^2	0.871	0.876	0.692	0.702

Notes: The dependent variable is the sales share of the largest establishments (top 5% or top 10% by sales) in each prefecture-year (columns 1–2) or industry-prefecture-year (columns 3–4). Working-age population growth is measured at the prefecture-year level. Standard errors are two-way clustered by prefecture and year (columns 1–2) or clustered at the prefecture-year level (columns 3–4). Columns (3)–(4) include only industry-prefecture-year cells with at least 100 establishments. Data from Census of Manufacturing Plants, 1985–2018. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B.8: Average population age and employment concentration

	Sales Share		Employment Share	
	(1) Top 5%	(2) Top 10%	(3) Top 5%	(4) Top 10%
Average age	−0.00723***	−0.00607***	−0.00484***	−0.00511***
	(0.00217)	(0.00201)	(0.00140)	(0.00131)
Industry × Prefecture FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	19,913	19,913	19,913	19,913
Adjusted R^2	0.692	0.702	0.733	0.733

Notes: The dependent variable is the sales share (columns 1–2) or employment share (columns 3–4) of the largest establishments (top 5% or top 10%) in each industry-prefecture-year cell. Average age is measured at the prefecture-year level. Standard errors are clustered at the prefecture-year level. All columns include only industry-prefecture-year cells with at least 100 establishments. Data from Census of Manufacturing Plants, 1985–2018. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

C Theoretical Foundations and Identification

This section provides formal proofs of propositions from the static model, efficiency characterizations under alternative production technologies, and identification arguments for the structural parameters.

C.1 Proofs of Propositions from the Static Model

This subsection provides rigorous proofs for all propositions stated in Section 3.1 of the main text. We establish conditions under which the planner's allocation dominates competitive equilibrium and characterize the optimal labor reallocation across worker groups.

Preliminary Lemmas

Lemma 1. *Let $\mathcal{P}$ be a partition of $\mathcal{G}$, and let $\{l_g^c(z)\}$ denote the competitive equilibrium allocation. Then $\{l_g^c(z)\}$ is feasible for the planner's problem under partition $\mathcal{P}$.*

Proof. In competitive equilibrium, market clearing requires $\int l_g^c(z)d\Phi(z) = L_g$ for all $g \in \mathcal{G}$.

For any partition element $P_s \in \mathcal{P}$, the planner's constraint is:

$$\int \sum_{g \in P_s} l_g(z)d\Phi(z) = L_s \equiv \sum_{g \in P_s} L_g.$$

Substituting the competitive allocation:

$$\int \sum_{g \in P_s} l_g^c(z)d\Phi(z) = \sum_{g \in P_s} \int l_g^c(z)d\Phi(z) = \sum_{g \in P_s} L_g = L_s.$$

Thus, $\{l_g^c(z)\}$ satisfies all partition constraints, establishing feasibility. □

Lemma 2. *If $\mathcal{P}$ is a refinement of $\mathcal{P}'$, then any allocation feasible under $\mathcal{P}$ is also feasible under $\mathcal{P}'$.*

Proof. Since $\mathcal{P}$ refines $\mathcal{P}'$, every element $P_s \in \mathcal{P}$ is contained in some element $P_t' \in \mathcal{P}'$. Consider an allocation $\{l_g(z)\}$ feasible under $\mathcal{P}$:

$$\int \sum_{g \in P_s} l_g(z)d\Phi(z) = L_s \quad \forall P_s \in \mathcal{P}.$$

For any $P'_t \in \mathcal{P}'$, let $\mathcal{S}_t = \{P_s \in \mathcal{P} : P_s \subseteq P'_t\}$. Since $\mathcal{P}$ partitions each P'_t:

$$\int \sum_{g \in P'_t} l_g(z) d\Phi(z) = \sum_{P_s \in \mathcal{S}_t} \int \sum_{g \in P_s} l_g(z) d\Phi(z) = \sum_{P_s \in \mathcal{S}_t} L_s = L'_t.$$

Hence, $\{l_g(z)\}$ is feasible under $\mathcal{P}'$. □

Proofs of Main Results

Proof of Proposition 1. We show that $Y^p(\mathcal{P}) \geq Y^c$ whenever $|\mathcal{P}| = S < G$.

Step 1: Weak inequality. By Lemma 1, the competitive allocation $\{l^c_g(z)\}$ is feasible for the planner. Since the planner maximizes over a set containing this allocation:

$$Y^p(\mathcal{P}) = \max_{\{l_g(z)\} \in \mathcal{F}(\mathcal{P})} \int F(\{l_g\}; z) d\Phi(z) \geq \int F(\{l^c_g\}; z) d\Phi(z) = Y^c,$$

where $\mathcal{F}(\mathcal{P})$ denotes the feasible set under partition $\mathcal{P}$.

Step 2: Strict inequality (under non-degeneracy). Suppose there exist $g, g' \in P_s$ for some s such that $w_g \neq w_{g'}$. In a competitive equilibrium, the first-order conditions imply:

$$\frac{\partial F(\{l^c_g\}; z)}{\partial l_g} = w_g \neq w_{g'} = \frac{\partial F(\{l^c_g\}; z)}{\partial l_{g'}} \quad \text{for all } z.$$

However, the planner's optimality conditions require:

$$\frac{\partial F(\{l^p_g\}; z)}{\partial l_g} = \mu_s = \frac{\partial F(\{l^p_g\}; z)}{\partial l_{g'}} \quad \text{for all } g, g' \in P_s.$$

Since marginal products are not equalized at the competitive allocation, while they must be equalized at the planner's optimum, we have $\{l^p_g(z)\} \neq \{l^c_g(z)\}$. By strict concavity of F, the optimum is unique, implying $Y^p(\mathcal{P}) > Y^c$. □

Proof of Proposition 2. Let $g, g' \in P_s$ with $w_g > w_{g'}$ in competitive equilibrium. We show that $\frac{L^p_g}{L^p_{g'}} > \frac{L^c_g}{L^c_{g'}} = \frac{L_g}{L_{g'}}$.

Step 1: Characterize marginal products at competitive equilibrium. In competitive equilibrium, the first-order conditions for firm optimization yield:

$$\frac{\partial F(\{l^c_g\}; z)}{\partial l_g} = w_g > w_{g'} = \frac{\partial F(\{l^c_g\}; z)}{\partial l_{g'}} \quad \text{for all } z.$$

Step 2: Show the competitive allocation is suboptimal for the planner. Consider a perturbation from the competitive allocation: for each z, increase $l_g(z)$ by ϵ and decrease $l_{g'}(z)$ by ϵ, where $\epsilon > 0$ is small. This perturbation satisfies the planner's constraint since aggregate labor in P_s is unchanged:

$$\int \left[(l_g^c(z) + \epsilon) + (l_{g'}^c(z) - \epsilon) + \sum_{h \in P_s \setminus \{g,g'\}} l_h^c(z) \right] d\Phi(z) = L_s.$$

The first-order change in output from this perturbation is:

$$\int \left[\frac{\partial F}{\partial l_g}(\{l_g^c\}; z) - \frac{\partial F}{\partial l_{g'}}(\{l_g^c\}; z) \right] \epsilon \, d\Phi(z) = (w_g - w_{g'})\epsilon > 0.$$

Thus, output strictly increases by shifting labor from group g' to group g.

Step 3: Establish the inequality. Since the competitive allocation is not optimal for the planner, and strict concavity ensures a unique optimum, the planner's optimal allocation must differ from the competitive allocation. By Step 2, the direction of improvement involves increasing L_g and decreasing $L_{g'}$ while satisfying $L_g^p + L_{g'}^p = L_g + L_{g'}$.

At the planner's optimum, we therefore have $L_g^p > L_g = L_g^c$ and $L_{g'}^p < L_{g'} = L_{g'}^c$. This implies:

$$\frac{L_g^p}{L_{g'}^p} > \frac{L_g}{L_{g'}} = \frac{L_g^c}{L_{g'}^c},$$

where the strict inequality follows because the numerator increases and the denominator decreases. □

Proof of Proposition 3. Let $\mathcal{P}$ be a refinement of $\mathcal{P}'$. Denote by $\{l_g^p(z; \mathcal{P})\}$ the planner's optimal allocation under partition $\mathcal{P}$.

By Lemma 2, $\{l_g^p(z; \mathcal{P})\}$ is feasible under $\mathcal{P}'$. Therefore:

$$Y^p(\mathcal{P}') = \max_{\{l_g(z)\} \in \mathcal{F}(\mathcal{P}')} \int F(\{l_g\}; z) d\Phi(z) \geq \int F(\{l_g^p(\cdot; \mathcal{P})\}; z) d\Phi(z) = Y^p(\mathcal{P}).$$

For strict inequality, suppose $\mathcal{P} \neq \mathcal{P}'$. Then there exists some $P_t' \in \mathcal{P}'$ that properly contains multiple elements of $\mathcal{P}$. If these elements have workers with different competitive wages, Proposition 1 implies $Y^p(\mathcal{P}') > Y^p(\mathcal{P})$. □

Proof of Corollary 1. Part 1: The maximal partition $\mathcal{P}^{max} = \{\mathcal{G}\}$ is coarser than any proper partition $\mathcal{P} \neq \mathcal{P}^{max}$, meaning $\mathcal{P}$ refines $\mathcal{P}^{max}$. By Proposition 3, $Y^p(\mathcal{P}^{max}) \geq Y^p(\mathcal{P})$.

Strict inequality holds if there exist groups $g, g' \in \mathcal{G}$ with $w_g \neq w_{g'}$ that belong to different elements of $\mathcal{P}$. Since $\mathcal{P}^{max}$ places all groups in a single element while $\mathcal{P}$ separates some groups, the planner gains additional flexibility under $\mathcal{P}^{max}$.

Part 2: Under $\mathcal{P}^{min} = \{\{g\} : g \in \mathcal{G}\}$, each partition element contains exactly one group. The planner's constraints become:

$$\int l_g(z) d\Phi(z) = L_g \quad \forall g \in \mathcal{G},$$

which are identical to the competitive market clearing conditions (2). Since both problems maximize the same objective subject to the same constraints, $Y^p(\mathcal{P}^{min}) = Y^c$. □

CES-Specific Results

Proof of Proposition 4. Under CES production, the marginal rate of technical substitution between groups g and g' at firm z is:

$$\text{MRTS}_{g,g'}(z) \equiv \frac{\partial F/\partial l_g}{\partial F/\partial l_{g'}} = \left(\frac{A_g(z)}{A_{g'}(z)}\right)^{1/\sigma} \left(\frac{l_{g'}(z)}{l_g(z)}\right)^{1/\sigma},$$

which is strictly decreasing in $l_g(z)/l_{g'}(z)$. In competitive equilibrium, profit maximization requires $\text{MRTS}_{g,g'}(z) = w_g/w_{g'} > 1$ for all z. For $g, g' \in P_s$, the planner's first-order conditions instead require $\text{MRTS}_{g,g'}(z) = 1$ for all z. Since the MRTS is strictly decreasing in $l_g(z)/l_{g'}(z)$ and must fall from $w_g/w_{g'} > 1$ to 1, we have $l_g^p(z)/l_{g'}^p(z) > l_g^c(z)/l_{g'}^c(z)$ for all z. □

C.2 Efficiency Analysis Under CES Production

This subsection analyzes how the efficiency gains from optimal labor reallocation depend on the elasticity of substitution between worker types and the dispersion of comparative advantages across firms.

Proposition 5. *Under the CES production technology, aggregate output under the planner's optimal allocation exceeds competitive equilibrium output if and only if $\sigma < \infty$ and there exists heterogeneity in the productivity ratios $A_g(z)/A_{g'}(z)$ across firms. The efficiency gain is increasing in the dispersion of comparative advantage across firm-worker pairs.*

Proof. When $\sigma \to \infty$, worker groups become perfect substitutes and the production function reduces

to:

$$y(z) = \left(\sum_g A_g(z) l_g(z) \right)^{\alpha}.$$

In this case, only total labor $\sum_g l_g(z)$ matters for production, making the composition of labor irrelevant conditional on total employment. Both the competitive equilibrium and the planner's solution yield identical output.

When $\sigma < \infty$, groups are imperfect substitutes. The planner can exploit firm-level comparative advantage by allocating relatively more workers of type g to firms where $A_g(z)$ is high relative to other groups. The competitive equilibrium cannot achieve this because each group's labor supply is fixed.

Formally, the efficiency gain can be expressed as:

$$\Delta Y \equiv Y^p - Y^c = \int \left[F(\{l_g^p\}; z) - F(\{l_g^c\}; z) \right] d\Phi(z).$$

By a second-order Taylor expansion around the competitive allocation and using the envelope theorem, this gain is proportional to the variance of productivity ratios $\text{Var}_{z,g}(A_g(z)/\bar{A}_g)$, where $\bar{A}_g = \int A_g(z) d\Phi(z)$. □

C.3 Analytical Results: Cobb-Douglas Case

This subsection derives closed-form solutions for the competitive equilibrium and social planner's problem under Cobb-Douglas production technology, providing analytical intuition for the efficiency gains from labor reallocation.

Suppose a simple example where $F(\cdot)$ is the sum of Cobb-Douglas functions in labor and firm-worker match productivity $A_g(z)$,

$$F(\{l_g\}; z) = \sum_g l_g^{\alpha} A_g(z)^{1-\alpha}.$$

Firm z solves,

$$\max_{l_g} \sum_g l_g^{\alpha} A_g(z)^{1-\alpha} - \sum_g w_g l_g$$

From the FOC,

$$l_g^c(z) = \left(\frac{\alpha}{w_g} \right)^{1/(1-\alpha)} A_g(z)$$

Let $\bar{A}_g := \int A_g(z) d\Phi(z)$. From the market-clearing conditions,

$$\int l_g^c(z)\ d\Phi(z) = L_g \implies w_g = \alpha \left(\frac{\bar{A}_g}{L_g} \right)^{1-\alpha}$$

So the labor demand for firm z is proportional to the productivity of the firm relative to the aggregate/average productivity,

$$l_g^c(z) = L_g \frac{A_g(z)}{\bar{A}_g}$$

Total firm labor is,

$$l^c(z) := \sum_g l_g^c(z) = \sum_g L_g \frac{A_g(z)}{\bar{A}_g}$$

Output for firm z is,

$$y^c(z) = \sum_g \left(\frac{L_g}{\bar{A}_g} \right)^{\alpha} A_g(z)$$

Aggregate output for the whole economy is,

$$Y^c := \int y^c(z) d\Phi(z) = \int \sum_g \left(\frac{L_g}{\bar{A}_g} \right)^{\alpha} A_g(z) d\Phi(z) = \sum_g L_g^{\alpha} \bar{A}_g^{1-\alpha} \tag{C.15}$$

Planner's Problem

Suppose the planner can substitute across any group, so the partition $\mathcal{P}$ corresponds to the entire $\mathcal{G}$. By Corollary 1 this gives the maximal gain from labor reallocation

The planner's problem is,

$$\max_{l_g(z) \geq 0} \int \sum_g l_g^{\alpha} A_g(z)^{1-\alpha} d\Phi(z) \quad s.t \quad \int \sum_g l_g(z) d\Phi(z) = L$$

From the FOCs,

$$[l_g(z)] : \ \alpha l_g^{\alpha-1} A_g(z)^{1-\alpha} = \lambda \quad \forall g, z$$

This implies marginal productivity is equalized across all worker groups and firm types,

$$\alpha l_g^{\alpha-1} A_g(z)^{1-\alpha} = \alpha l_{g'}^{\alpha-1} A_{g'}(z')^{1-\alpha} \implies l_{g'}(z') = \frac{A_{g'}(z')}{A_g(z)} l_g(z) \ \forall g, g', z, z'.$$

Substituting in the resource constraint,

$$\begin{aligned} \sum_{g'} l_{g'}(z') d\Phi(z') &= L \\ l_g(z) \sum_{g'} \int \frac{A_{g'}(z')}{A_g(z)} d\Phi(z') &= L \\ \implies l_g^p(z) &= \frac{A_g(z)}{\bar{A}} L, \end{aligned}$$

where $\bar{A} \coloneqq \int \sum_{g'} A_{g'}(z') d\Phi(z')$

Aggregating across firms implies that the optimal labor supply allocation is,

$$l_g^p = \frac{\bar{A}_g}{\bar{A}} L.$$

The total aggregate output is,

$$Y^p = \int \sum_g l_g^p(z)^\alpha A_g(z)^{1-\alpha} d\Phi(z) = L^\alpha \bar{A}^{1-\alpha}$$

So the maximal output gain from labor supply reallocation is,

$$\Delta\%Y \coloneqq \frac{Y^p}{Y^c} - 1 = \frac{L^\alpha \bar{A}^{1-\alpha}}{\sum_g L_g^\alpha \bar{A}_g^{1-\alpha}} - 1$$

Notice the gain is non-monotonic in α as $\alpha = 0$ or $\alpha = 1$, implies $\Delta\%Y = 0$.

C.4 Extension: Generalized Production Function

This subsection extends the baseline production function to allow for more flexible complementarities between firm productivity, worker skill, and experience. The generalized specification accommodates non-monotonic age-productivity profiles.

We can make the production function more general as

$$y(z, h_i, x_j, l_{ij}) = \int_{h_i \in H, x_j \in X} \left[l_{ij}^{\frac{\sigma-1}{\sigma}} f(z, h_i, x_j; \nu_h, \nu_x) \right] dh_i dx_j, \tag{C.16}$$

where the group specific output $f_{ij} := f(z, h_i, x_j; \nu_h, \nu_x)$ has to satisfy the property that it is log super-modular w.r.t z, h_i and log sub-modular w.r.t z, x_j. In particular, parameters ν_h (and ν_x) govern the complementary between z and h_i (and z and x_j) respectively. We need $\nu_h > \nu_a$. Practically, the production function can include polynomial terms of h_i and x_j:

$$y(z, h_i, x_j, l_{ij}) = \int_{h_i \in H, x_j \in X} \left[l_{ij}^{\frac{\sigma-1}{\sigma}} \left(z^{\nu_h}(b_0 + b_1 h_i + b_2 h_i^2) + z^{\nu_a}(d_0 + d_1 x_j + d_2 x_j^2) \right) \right] dh_i dx_j, \quad \text{(C.17)}$$

where d_2 can be negative, meaning that experience (age) contributes to output negatively when it exceeds a threshold.

C.5 Identification Strategy and Theory

This subsection formalizes the identification of structural parameters from observed moments. We demonstrate that the employment distribution and wage profiles uniquely determine the match-specific productivity parameters under the baseline Cobb-Douglas specification.

Suppose the case where there are two types of firms $k \in \{s, l\}$ and two skills $i \in \{1, 2\}$ and $\sigma = \frac{1}{1-\alpha}$ so we are in the Cobb-Douglas case.

From the firm optimization problem and the market-clearing condition,

$$l_{ik} = L_i \frac{A_{ik}}{\bar{A}_i} \ , \quad w_i = \alpha \left(\frac{\bar{A}_i}{L_i} \right)^{1-\alpha},$$

where $\bar{A}_i = \sum_k A_{ik}$.

Summing up across skills i and using the equilibrium labor at the firm level L_k,

$$L_k = \sum_i L_i \frac{A_{ik}}{\bar{A}_i}.$$

Now we can also sum up the wages at the firm level,

$$w_k = \sum_i w_i \frac{L_{ik}}{L_k} = \sum_i \alpha \left(\frac{\bar{A}_i}{L_i} \right)^{1-\alpha} \frac{L_i}{L_k} \frac{A_{ik}}{\bar{A}_i} = \sum_i \alpha \left(\frac{\bar{A}_i}{L_i} \right)^{-\alpha} \frac{A_{ik}}{L_k}$$

Consider the system of equations for $(A_{1s}, A_{1l}, A_{2s}, A_{2l}) \in \mathbb{R}^4_+$:

$$\frac{\bar{A}_1}{L_1} A_{1s} + \frac{\bar{A}_2}{L_2} A_{2s} = L_s \tag{C.18}$$

$$\frac{\bar{A}_1}{L_1} A_{1l} + \frac{\bar{A}_2}{L_2} A_{2l} = L_l \tag{C.19}$$

$$\left(\frac{\bar{A}_1}{L_1}\right)^{-\alpha} A_{1s} + \left(\frac{\bar{A}_2}{L_2}\right)^{-\alpha} A_{2s} = \frac{w_s}{\alpha} L_s \tag{C.20}$$

$$\left(\frac{\bar{A}_1}{L_1}\right)^{-\alpha} A_{1l} + \left(\frac{\bar{A}_2}{L_2}\right)^{-\alpha} A_{2l} = \frac{w_l}{\alpha} L_l \tag{C.21}$$

where $\bar{A}_1 := A_{1s} + A_{1l}$, $\bar{A}_2 := A_{2s} + A_{2l}$, and parameters satisfy $\alpha \in (0,1)$, $L_1, L_2 > 0$.

The system (C.18)–(C.21) reduces to a two-dimensional system for aggregate allocations. Adding equations (C.18) and (C.19):

$$\frac{(A_{1s} + A_{1l})\bar{A}_1}{L_1} + \frac{(A_{2s} + A_{2l})\bar{A}_2}{L_2} = L_s + L_l$$

Since $\bar{A}_i = A_{is} + A_{il}$, this yields:

$$\frac{\bar{A}_1^2}{L_1} + \frac{\bar{A}_2^2}{L_2} = L_s + L_l \tag{C.22}$$

Similarly, adding (C.20) and (C.21):

$$\bar{A}_1 \left(\frac{\bar{A}_1}{L_1}\right)^{-\alpha} + \bar{A}_2 \left(\frac{\bar{A}_2}{L_2}\right)^{-\alpha} = \frac{w_s L_s + w_l L_l}{\alpha}$$

which simplifies to:

$$\bar{A}_1^{1-\alpha} L_1^{\alpha} + \bar{A}_2^{1-\alpha} L_2^{\alpha} = \frac{w_s L_s + w_l L_l}{\alpha} \tag{C.23}$$

For data values of w_i and L_i, we can use C.22 and C.23 to solve for $\bar{A}_i$.

C.6 Computational Algorithm for Dynamic Equilibrium

This subsection describes the numerical algorithm used to solve for the dynamic competitive equilibrium. The key challenge is that firms' current employment decisions depend on expected future firing costs, which themselves depend on future employment choices.

In this section, we describe the algorithm we use to solve the dynamic model from Section 3.2.

In our setting, two special experience bins do not require employment information in two periods

when solving for the current employment choice. First, the firm does not pay the firing cost when choosing employment of workers with the least experience, as there is no employment stock for this group of workers. Second, for bins with the most experience (i.e., the oldest workers), the firm does not need to take into account the firing cost it would have to pay next period as they all retire by the end of the period. Therefore, we can solve the employment choices for the youngest workers first and proceed to solve the employment choices *sequentially* from experience group x_j to experience group x_{j+1} where $j \in [1, J-1]$. Other than the wage schedule, key equilibrium objects needed for solving the employment choices are the expected firing probabilities of a worker with h_i and x_j hired this period (i.e., $\mathbb{E}\left[\mathbb{1}[\tilde{L}_{i,j+1} > l_{i,j+1}]\right]$ and $\mathbb{E}\left[\mathbb{1}[\tilde{L}_{i+1,j+1} > l_{i+1,j+1}]\right]$).

Practically, we solve the labor market equilibrium using the following steps.

1. First, we solve the model without the firing cost and obtain the wage schedule $w^{no}_{i,j}$ and firms' employment choices, $l^{no}(z)_{ij}$ ($j \in [1,2,3,...,J]$), and employment, stocks $\tilde{L}^{no}_{ij}$ ($j \in [2,3,...,J]$). Based on these solved equilibrium objects, we construct the expected firing costs that show up in the right hand side of equations (9) and (10) denoted by FC^0_{ij}.

2. Next, we solve for $w_{i,1}$ that equalizes the labor supply in the submarket ($g(h_i, x_1)$) to the aggregated labor demand across all active firms whose firm-level labor demand is denoted by $l_{i,1}(z)$. For this step, we assume each firm uses the expected firing cost constructed from the previous step. Having solved $w_{i,1}$, we calculate firms' employment choices of the youngest workers, $l_{i,1}(z)$, and their employment stocks of the second youngest experience bins $\tilde{L}_{i,2}(z)$.

3. Third, we solve for $w_{i,j}$ (where $j \in [2,T]$) that equalizes the labor supply in the submarket ($g(h_i, x_j)$) to the aggregated labor demand across all active firms *sequentially* (i.e., from experience group x_{j-1} to experience group x_j). Importantly, we use the *derived* $\tilde{L}_{i,j}(z)$ calculated from the employment choice made for the experience group x_{j-1}. For this step, we also assume each firm uses the expected firing cost constructed from the first step. Note that when $j = T$, there is no need to calculate the expected firing cost as all workers retire at the end of experience x_J.

Using the above three steps, we derive the wage schedule, $w_{i,j}$, employment choices and stocks, $l_{i,j}(z)$ and $\tilde{L}_{i,j}(z)$, which lead to the newly calculated values of the expected firing cost:

$$\begin{aligned} FC^1_{ij}(z): \quad = \quad & \eta\tau\delta\Bigg[(1-p_i-q_i+2p_iq_i)w_{ij+1}\mathbb{E}\left[\mathbb{1}[\tilde{L}_{i,j+1} > l_{i,j+1}]\right] \\ & +p_i(1-q_i)w_{i+1,j+1}\mathbb{E}\left[\mathbb{1}[\tilde{L}_{i+1,j+1} > l_{i+1,j+1}]\right] + q_i(1-p_i)w_{i-1,j+1}\mathbb{E}\left[\mathbb{1}[\tilde{L}_{i-1,j+1} > l_{i-1,j+1}]\right]\Bigg], \end{aligned}$$

where $w_{i,j}$, $l_{i,j}(z)$ and $\tilde{L}_{i,j}(z)$ are derived in the second and third steps. In the final step, we update the expected firing cost by taking a weighted average of the assumed expected firing cost in previous steps and newly calculated expected firing cost:

$$FC_{ij}^2 = \zeta FC_{ij}^1 + (1-\zeta)FC_{ij}^0,$$

where $\zeta \in (0,1)$ is the adjustment factor. We then iterate our procedure by using FC_{ij}^2 to replace FC_{ij}^0 and redoing steps two to four. We stop until the newly calculated wage schedule and employment choices are the same as the ones calculated from the previous iteration.[32]

[32] In such a case, the newly calculated expected firing cost is also the same as the one calculated from the previous iteration.