The Worker and Firm Origins of Life-Cycle Wage Inequality*

Xincheng Qiu† Peking University

Jesse Wedewer‡ Duke

Runling Wu§ Duke

March 2026

(click here for latest version)

Abstract

This paper studies how worker and firm heterogeneity shape wage inequality over the life cycle. We extend the Abowd, Kramarz and Margolis (1999) (AKM) framework to allow for worker and firm heterogeneity in wage growth, and develop a two-way clustering algorithm to estimate it. Using matched employer-employee data from Italy, we document substantial heterogeneity in both worker and firm growth components. Decomposing life-cycle wage inequality reveals that differences in worker growth component account for most of the increase, with firm heterogeneity playing a smaller role. Positive sorting between high-growth workers and high-wage firms intensifies over the life cycle, amplifying inequality through both steeper growth trajectories and differential access to premium employers.

Keywords: Two-sided heterogeneity, wage growth, life-cycle wage inequality
JEL Codes: J00, J31, J40

*First draft: December, 2024. We thank Peter Arcidiacono, Patrick Bayer, Vince Chen, Michael Dinerstein, Rafael Dix-Carneiro, Victoria Gregory, Daniel Hannwinckel, Gregor Jarosch, Philipp Kircher, Arnaud Maurel, and Laura Pilossoph for helpful comments. We also thank seminar audiences at CityU, Duke, Cleveland Fed, HKU, Minneapolis Fed, and SED Copenhagen. All remaining errors are our own.

†Email: xincheng.qiu@gsm.pku.edu.cn.
‡Email: jesse.wedewer@duke.edu.
§Email: runling.wu@duke.edu.

1 Introduction

This paper studies how worker and firm heterogeneity shape the evolution of wage inequality over the life cycle of workers. A large literature has documented that wage inequality increases substantially over the life cycle. Some workers experience rapid and consistent wage growth over the course of their careers, whereas other workers have flat wage profiles. This divergence in wage-age profiles across workers may reflect three distinct mechanisms. First, workers may differ in their ability to accumulate human capital or their willingness to bargain over wages with their employer, leading to heterogeneous returns to experience across workers irrespective of their employer (Rubinstein and Weiss, 2006; Huggett et al., 2011; Bagger et al., 2014; Guvenen et al., 2021; Ozkan et al., 2023). Second, firms may differ in their organizational structures and the environments they provide for training or promotion opportunities, resulting in heterogeneity in wage growth opportunities across employers (Gregory, 2020; Arellano-Bover and Saltiel, 2024; Borovičková and Macaluso, 2024; Ma et al., 2024). Third, through on-the-job search, the sorting of workers to firms may evolve over the life cycle, with some workers climbing the job ladder faster than others (Burdett and Mortensen, 1998; Postel-Vinay and Robin, 2002). Our paper quantifies the relative importance of these channels, measuring the respective contribution of worker heterogeneity in returns to experience, firm heterogeneity in returns to experience, and the dynamic worker-firm sorting to life-cycle wage inequality. Applying our framework to the Veneto region of Italy, we find that differences in returns to experience across workers account for nearly all of the increase in wage dispersion over the life-cycle, with firm heterogeneity and worker-firm sorting playing smaller roles.

To quantify these channels, we propose a statistical model that extends the Abowd, Kramarz and Margolis (1999) (hereafter, AKM) framework, allowing for returns to experience to vary across workers and employers. As in the canonical framework, wages depend on a worker component and a firm wage premium, common to all workers. In our model, the worker component is dynamic, comprising a static worker effect and a dynamic growth effect. The worker effect captures differences in pay among early-career workers at the same firm, while the growth effect reflects cumulative returns to experience, which are general and portable across jobs. The growth effect captures two key sources of heterogeneity. First, returns to experience, *worker growth component*, differ across workers irrespective of where they acquire the experience. Second, returns to experience are allowed to vary across employers, *firm growth component*, allowing experience at each employer to influence wages differently. Under a conditional random mobility assumption, we separately identify the worker growth component from the firm growth component, exploiting variations in the composition of experience acquired at different firms over the course of workers' careers as they move from employer to employer.

To implement this framework, we develop a two-way clustering algorithm that extends the

group fixed-effect approach of Bonhomme and Manresa (2015) to two-sided heterogeneity, restricting heterogeneity in worker experience effects and firm experience effects to a discrete number of worker and firm growth types. Our approach minimizes the sum of squared residuals between observed wages and predicted wages from our extended AKM specification, which includes worker fixed effects, firm fixed effects, worker growth components, and firm growth components. The key challenge is to jointly determine the assignment of workers and firms to growth types while estimating group-specific returns to experience.

Our algorithm overcomes this challenge through an iterative search procedure, alternating between updating parameters and reassigning units to groups. The algorithm leverages a key insight: conditional on group assignments, parameter estimation reduces to standard Ordinary Least Squares (OLS) regression, while conditional on parameters, optimal assignments can be determined independently for each unit. Starting from an initial guess of group assignment obtained through k-means clustering, the algorithm alternates between three steps until the assignments and parameters converge.[1] First, given the current group assignments, it updates all model parameters through OLS regression. Second, holding firm assignments and parameters fixed, it reassigns workers to groups to minimize the sum of squared residuals. Third, holding worker assignments and parameters fixed, it reassigns firms to groups to minimize the sum of squared residuals. To mitigate dependence on initialization, we run the algorithm from multiple starting points and retain the solution with the lowest objective function value.

We validate the performance of the algorithm through extensive Monte Carlo simulations. First, we examine the performance under scenarios with and without worker-firm sorting, demonstrating that it articulately recovers the relative importance of worker versus firm contributions to wage growth. Second, we systemically validate the algorithm under alternative data-generating processes where worker effects dominate, where firm effects dominate, and where both sources contribute equally to wage growth, confirming that the algorithm correctly identifies the relative importance of each channel. Lastly, we establish robustness across alternative cluster-pair specifications, demonstrating that our baseline specification of 20 groups on both sides captures the key heterogeneity in growth patterns while maintaining computational tractability.[2]

We apply our statistical model and estimation framework to understand wage dynamics of young workers in the Veneto region of Italy from 1984 through 2001. Our analysis yields three primary findings on the sources of life-cycle wage inequality.

[1] Our initial guess of group assignments is from a k-means clustering, where workers are grouped based on the empirical distribution of wages over the course of their careers and firms are grouped based on the distribution of annual wage changes among job stayers.

[2] We experiment by keeping the number of worker groups fixed while varying the number of firm groups between 20 and 50, and vice versa, to ensure our results are not driven by our particular choice of group numbers.

First, differences in worker growth components account for most of the increase in life-cycle wage inequality. Our analysis reveals substantial heterogeneity across worker growth types. While workers in the majority of types exhibit relatively flat wage growth profiles, workers in high-growth types experience wage growth exceeding 4 percent per year of experience.[3] We decompose the variance of wages for a cohort of male workers born in 1959 over the course of their careers. The increasing dispersion of life-cycle wage growth is primarily driven by heterogeneity in the worker growth component rather than the firm growth component. As workers accumulate more experience over their careers, those belonging to high-growth types diverge sharply from those in flat-growth types, progressively widening the life-cycle wage dispersion. In contrast, the firm growth component contributes relatively little to rising life-cycle inequality due to two key features of the Veneto labor market: substantially lower dispersion across firm growth types and high worker mobility across firms, which limits the cumulative impact of employment at any particular firm type.

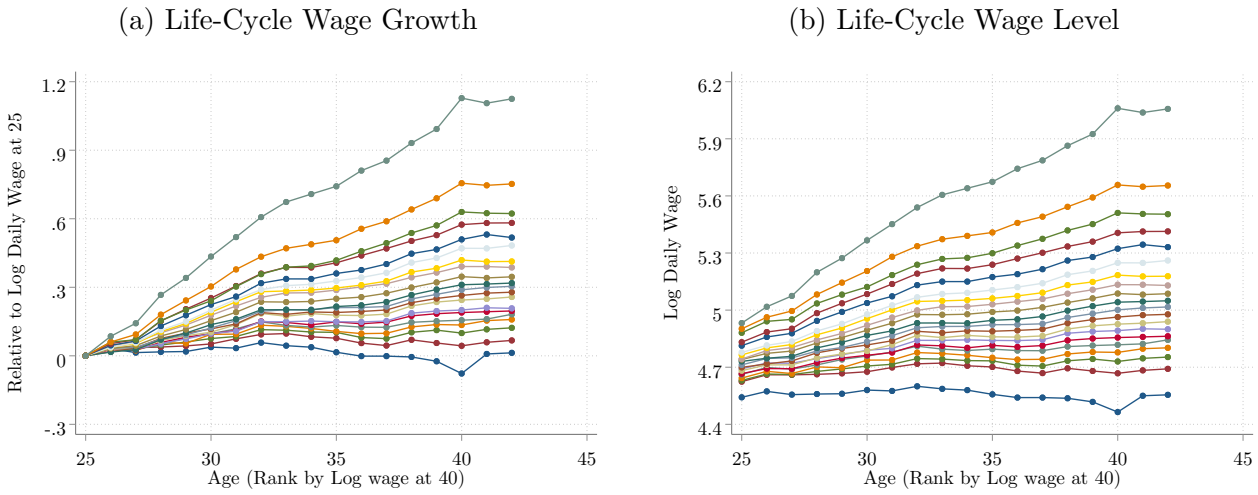
Second, we document sorting between high-growth workers and high-wage firms that strengthens over the life-cycle. A large literature has established that high-wage workers (those with high worker fixed effects) tend to match with high-wage firms (those with high firm fixed effects) (Abowd et al., 1999; Card et al., 2013, 2016; Song et al., 2019). Our framework extends this understanding by decomposing the dynamic worker effect into two components: a worker fixed effect capturing initial wage levels and a worker growth component capturing the steepness of wage trajectories over the life cycle. We find that the covariance between the dynamic worker effect and firm fixed effects increases substantially as workers age. Early in their careers, workers with high growth components already tend to match with firms offering higher firm wage premia, but this initial sorting is relatively modest. As workers accumulate experience, however, this association intensifies considerably. By their early 40s, high-growth workers have become increasingly concentrated at firms that pay all their employees more generously. This strengthening sorting over the life cycle reveals the dynamic process underlying positive assortative matching—workers with steeper wage growth trajectories progressively sort into higher-paying firms, amplifying wage inequality through both their individual growth advantage and their access to premium employers.[4]

Figure 1 illustrates this stark divergence in life-cycle wage profiles. Tracking male workers in Veneto from age 25 onward, the figure reveals dramatic differences in wage trajectories across 20 groups of workers, which are grouped based on their log wage at age 40. Panel (a) shows life-cycle wage growth, the difference between log daily wages at each age and log daily wages

[3]Our methodology groups workers into twenty distinct growth types, allowing us to characterize heterogeneity through group-specific growth components rather than individual-specific parameters.

[4]Consistent with Arellano-Bover and Saltiel (2024), we fail to find meaningful relationships between high growth employers and other components determining pay. Employers offering high returns to experience are no more likely to have large firm wage premia.

Figure 1: Life-Cycle Wage Profile



Notes: This figure compares life-cycle wage trajectories of workers in Veneto, classified into 20 groups based on their wage rank at age 40. Panel (a) displays life-cycle wage growth, measured as the log daily wage at each age minus the log daily wage at age 25. Panel (b) presents the actual log daily wage levels over the life cycle. Both panels use a sample of male workers.

at age 25. While all workers start at zero, their wage growth trajectories diverge substantially over time. Workers in the top wage group at age 40 have experienced cumulative wage growth exceeding 120 percent, whereas those in the bottom see virtually no real wage growth over the same fifteen-year period. Panel (b) displays the actual log wage levels, revealing that these differential growth patterns translate into substantial and widening lifetime earnings inequality across workers.

Related literature. We contribute to the growing literature that considers the role of firms in the determination of labor market outcomes using high-dimensional fixed effect models to control for unobserved worker heterogeneity (Abowd et al., 1999; Card et al., 2013, 2016; Song et al., 2019). We build on this literature by considering the role of firm and worker heterogeneity in the determination of the returns to experience, extending the framework of Abowd et al. (1999) to admit two-sided heterogeneity in wage-experience profiles of workers. In the spirit of Bonhomme and Manresa (2015) and similar to a number of recent papers[5], we use a group fixed approach, modeling workers and firms as discrete types. Our key methodological contribution is a novel two-way clustering algorithm that jointly classifies workers and firms into growth types and estimates group-specific returns-to-experience coefficients through an iterative procedure. This distinguishes our approach from Arellano-Bover and Saltiel (2024), who use k-means clus-

[5]The number of papers using a group fixed approach in settings with two-sided heterogeneity is rapidly expanding. Some prominent examples include Dauth et al. (2022); Bradley et al. (2023); Gregory et al. (2024); Arellano-Bover and Saltiel (2024).

tering to group firms but restrict worker heterogeneity to initial wage levels through worker fixed effects. Our two-way approach enables us to quantify sorting between worker and firm growth types and decompose the sources of life-cycle wage inequality into contributions from worker heterogeneity, firm heterogeneity, and their interaction.

We also contribute to the large literature that considers the post-schooling accumulation of human capital. Similar to papers that allow for returns to experience to differ with educational attainment (Gathmann and Schönberg, 2010; Bagger et al., 2014; Helm et al., 2023; Deming, 2023), we allow for returns to experience to be heterogeneous across workers. We build on this work by allowing for unrestricted heterogeneity in returns to experience across workers. Consistent with evidence on heterogeneous income profiles in the earnings dynamics literature (Guvenen, 2009; Magnac et al., 2018), we find that there are large and systematic differences in returns to experience across workers. We incorporate this worker heterogeneity in a Mincerian regression framework that allows for returns to experience to vary across employers. In doing so, we build upon evidence from Sørensen and Vejlin (2011) and Gregory (2020), who consider two-sided heterogeneity in wage growth using a fixed effects approach. Similar to these papers, we find that there is substantially more dispersion in returns to experience across worker types than there is across firm types.

We use our framework to contribute to the literature that considers the evolution of wage inequality over the life-cycle of workers (Rubinstein and Weiss, 2006; Huggett et al., 2011; Bagger et al., 2014; Gregory, 2020; Guvenen et al., 2021; Ozkan et al., 2023; Arellano-Bover and Saltiel, 2024; Borovičková and Macaluso, 2024). We use our framework to decompose the sources of wage dispersion over the life-cycle of workers. We consider the respective roles of heterogeneity in returns to experience across workers and across firms. We find that heterogeneity in returns to experience across workers is the primary driver in explaining why wage dispersion increases over the life-cycle of workers, with heterogeneity in returns to experience across firms and the evolution of firm rents being relatively less important. In doing so, our findings complement recent work using frictional models of the labor market to understand the respective roles of worker and firm heterogeneity in wage growth (Gregory, 2020; Borovičková and Macaluso, 2024). While the methods and empirical contexts differ, our quantitative findings are broadly similar.

Roadmap. Section 2 presents our conceptual and empirical framework, extending the AKM model to allow for two-sided heterogeneity in returns to experience. Section 3 develops our two-way clustering algorithm. Section 4 validates it through Monte Carlo simulations and robustness checks. Section 5 describes the Veneto Worker History data. Section 6 presents baseline estimates and assesses the performance of the algorithm using individual-level fixed effects. Section 7 analyzes the sources of life-cycle wage inequality and dynamic sorting patterns.

Section 8 concludes.

2 Conceptual and Empirical Framework

This section introduces our conceptual and empirical framework for estimating heterogeneous returns to experience across workers and firms. We first present a conceptual model extending the AKM framework to allow for two-sided heterogeneity in wage growth, then develop our econometric specification using a group fixed effects approach that approximates the continuous distribution of returns to experience through discrete worker and firm types.

2.1 Conceptual Framework

Wage Determination. In the spirit of Abowd et al. (1999), we consider a statistical wage model where wages are an additively separable function of worker and firm identities. Formally, a worker i first entering the labor market at time t with characteristics $X_{i,t}$ and who secures employment at firm j earns a log wage according to:

$$\ln y_{i,t} = \underbrace{\alpha_i}_{\text{Worker effect}} + \underbrace{\psi_{j(i,t)}}_{\text{Firm effect}} + \underbrace{\xi X_{i,t}}_{\text{Observable factors}} + \underbrace{r_{i,t}}_{\text{Error term}} \tag{1}$$

The worker effect α_i reflects unobserved attributes possessed by workers upon entry to the labor market and that are rewarded equally across firms, whereas the firm effect, $\psi_{j(i,t)}$, reflects a pay premium offered to any worker at any stage of their career. Observed characteristics of workers or firms that shape wages are embedded in $\xi X_{i,t}$, with the error term $r_{i,t}$ capturing remaining wage variation.

As workers accumulate experience in the labor market, their wages evolve. We model the evolution of wages with experience *within* and *across* jobs through the flexible variable $\mathcal{H}_{i,t}$, which captures the cumulative returns to experience acquired by a worker over their employment history:

$$\ln y_{i,t} = \alpha_i + \psi_{j(i,t)} + \xi X_{i,t} + \underbrace{\mathcal{H}_{i,t}}_{\text{Growth effect}} + r_{i,t} \tag{2}$$

Heterogeneity in Returns to Experience. We assume that $\mathcal{H}_{i,t}$ is the sum of a worker growth component and a firm growth component:

$$\mathcal{H}_{i,t} = \underbrace{\beta_i \operatorname{Exp}_{i,t}}_{\text{worker growth component}} + \underbrace{\sum_{j' \in \mathcal{J}} \phi_j \mathbb{1}[j = j'] \mathbf{Exp^j_{i,t}}}_{\text{firm growth component}} \tag{3}$$

The *worker growth component* captures worker heterogeneity in cumulative returns to experience, which may arise either due to differences in time spent in the labor market or due to differences in how experience translates to wage growth through the worker experience effect β_i. We interpret β_i as embodying a few key dimensions of worker heterogeneity that influence wage growth similarly at *any* job. First, workers may vary systematically in their ability or willingness to bargain with employers over wages (Biasi and Sarsons, 2021; Roussille, 2024). Second, workers may differ in their capacity to accumulate human capital from experience or, alternatively, their willingness to invest in skill formation (Rubinstein and Weiss, 2006; Huggett et al., 2011). Third, workers may be heterogeneous in latent ability, which employers learn about as they acquire experience (Altonji and Pierret, 2001).

The *firm growth component* relates the heterogeneous employment histories of workers at different employers to wage growth. Experience attained at firm j, denoted as $\mathbf{Exp_{i,t}^{j}}$, influences wage rates according to the firm experience effect ϕ_j. Firm experience effects embody a few key dimensions of employer heterogeneity that influence wage growth of workers in similar ways, regardless of their characteristics. First, firms may vary in their responsiveness to outside offers and openness to wage renegotiation (Caldwell et al., 2024). Second, firms may differ in their organizational structures and opportunities they provide for training or promotion (Gregory, 2020; Ma et al., 2022; Arellano-Bover and Saltiel, 2024).

In contrast to the canonical AKM model, our framework generates heterogeneous wage profiles across workers through worker-specific returns to experience, and heterogeneous wage profiles across firms through firm-specific returns to experience. Moreover, the wage profiles in our model are history-dependent, shaped by workers' entire employment history including both their current employer and the sequence of firms at which they have previously worked.

2.2 Empirical Framework

We now present our econometric framework for disentangling the worker and firm heterogeneity in wage growth. In a given time interval, our data contains N^* person-year observations on N workers and J firms. The function $j(i,t)$ gives the identity of the firm that employs worker i in year t.

2.2.1 Two-Way Individual Fixed Effect

Our conceptual framework delivers a two-way individual fixed-effect regression specification. We assume that the log daily real wage $\ln y_{it}$ of individual i in year t is the sum of a worker component α_i, an firm component $\psi_{j(i,t)}$, a worker-specific returns to experience $\beta_i \operatorname{Exp}_{i,t}$, firm-specific returns to experience $\sum_j \phi_j \mathbf{Exp_{i,t}^{j}}$, an index of time-varying observable characteristics

$\xi X_{i,t}$, and an error component r_{it}:

$$\ln y_{i,t} = \alpha_i + \psi_{j(i,t)} + \beta_i \, \text{Exp}_{i,t} + \sum_j \phi_j \mathbf{Exp}^{\mathbf{j}}_{\mathbf{i,t}} + \xi X_{i,t} + r_{i,t} \tag{4}$$

where $X_{i,t}$ is a combination of life-cycle and aggregate factors (age, gender, and year) that affect worker i 's productivity at all jobs and $r_{i,t}$ is a mean-zero error term.[6]

While the individual fixed effects specification is conceptually appealing, it presents several practical limitations for our analysis. First, while we can directly estimate this specification, the coefficient estimates of individual fixed effects become unreliable and noisy due to high dimensionality, particularly for workers or firms with limited observations. The individual fixed effects approach relies on observation-by-observation estimation, i.e., i-by-i estimation, without exploiting information across observations, and consequently, we can learn only limited information from these noisy estimates. Second, a central goal of this paper is to quantify the relative importance of firm and worker heterogeneity in the evolution of life-cycle wage inequality through variance decomposition. However, existing variance correction methods fail in this context. Applying Kline et al. (2020) (hereafter, KSS) correction to individual-level interactive fixed effects encounters a fundamental rank deficiency problem. The KSS correction requires collapsing the data to the match level to address serial correlation within matches, but this procedure creates a system with more parameters than observations at the match level, i.e., more columns than rows, resulting in infinitely many possible solutions to the OLS problem and causing standard estimation methods, including conjugate gradient methods, to fail. Third, an alternative approach that estimates individual-level fixed effects using first-differences for job stayers exists and is more tractable, but it is restricted to workers who remain at the same employer for two consecutive years. This alternative approach fundamentally changes the sample composition by excluding short-tenure jobs, again limiting what we can learn from this exercise due to sample selection.[7]

2.2.2 Two-Way Group Fixed Effect

To address these challenges, we adopt a group fixed effects approach that approximates the underlying continuous distribution of individual fixed effects β_i, ϕ_J as suggested in Bonhomme and Manresa (2015), Bonhomme et al. (2022), and Bonhomme and Denis (2024). The key insight is to limit the number of values that β_i and ϕ_J can take, effectively reducing estimation

[6]Following AKM, we assume that the error term $r_{i,t}$ consists of three separate random effects: a match component $\eta_{i,j(i,t)}$, a unit root component $v_{i,t}$ and a transitory error $\varepsilon_{i,t}$: $r_{i,t} = \eta_{i,j(i,t)} + v_{i,t} + \varepsilon_{i,t}$.

[7]We discuss this alternative estimation procedure under the individual fixed effects framework and use it to validate our two-way clustering algorithm in Section 6. We present evidence comparing individual fixed effects and group fixed effects estimates in Appendix II.

noise by pooling workers and firms with similar wage growth characteristics while maintaining some flexibility to capture substantial heterogeneity in returns to experience on both sides.

Suppose that individual workers are partitioned into K_w groups, $\ell \in \{1, ..., K_w\}$, and that within group ℓ, all β_i's are constant and equal to β_ℓ. Similarly, suppose that firms are partitioned into K_f groups, $k \in \{1, ..., K_f\}$, and that within group k, all ϕ_j's are constant and equal to ϕ_k, with K_w and K_f denoting the number of worker and firm groups, respectively. We then consider the following two-way grouped fixed-effects model:

$$\ln y_{i,t} = \alpha_i + \psi_j + \underbrace{\beta_{\rho(i)} \operatorname{Exp}_{i,t}}_{\text{worker growth component}} + \underbrace{\sum_{\kappa(j)} \phi_{\kappa(j)} \mathbf{Exp}_{i,t}^{\kappa(j)}}_{\text{firm growth component}} + \xi X_{i,t} + r_{i,t} \tag{5}$$

where we have a worker assignment function $\rho(\cdot) : \{1, \dots I\} \to \{1, \dots, K_w\}$ that maps individual workers to worker groups, and a firm assignment function $\kappa(\cdot) : \{1, \dots, J\} \to \{1, \dots, K_f\}$ that maps individual firms to firm groups. $\beta_{\rho(i)}$ reflects heterogeneous returns to experience across worker groups, while $\phi_{\kappa(j)}$ reflects variation in returns to experience across firm groups. We refer to $\beta_{\rho(i)}$ and $\phi_{\kappa(j)}$ as the worker and firm experience effects, respectively, since they measure how one-year of working experience translates into earnings growth within each group. The full terms $\beta_{\rho(i)} \operatorname{Exp} i, t$ and $\sum \kappa(j) \phi_{\kappa(j)} \mathbf{Exp}_{i,t}^{\kappa(j)}$ are therefore referred as the worker growth component and firm growth component, representing the total contribution of worker and firm heterogeneity in returns to experience. Finally, we impose a normalization for firm group g^* so that $\phi_{g^*} = 0$, a restriction similar to the normalization requirement in Abowd et al. (1999).

We estimate equation (5) by OLS. For these estimates to be unbiased, the unobserved determinants of earnings must be uncorrelated with experience stocks, conditional on the worker's identity, their observable characteristics, and their contemporaneous employer. We assume that satisfies the strict exogeneity assumption:

$$\mathbb{E}[r_{i,t} | i, t, j(i,t), V_{i,t} = \{\operatorname{Exp}_{i,t}, \mathbf{Exp}_{i,t}^{\kappa(j)}, X_{i,t}\}] = 0 \tag{6}$$

where $V_{i,t}$ represents the set of experience measures and time-variant controls.

As shown by AKM and Abowd et al. (2002), the firm and person effects are only separately identified within a "connected set" of firms that are linked by worker mobility. In our analysis, we limit attention to workers and firms in the largest connected set.[8]

[8]Note that even within these connected sets, we require a normalization restriction on the worker and firm effects, as wage premia are only identified relative to a reference group.

3 Two-Way Clustering Algorithm

This section introduces our two-way clustering algorithm, which jointly classifies workers and firms into discrete growth types to estimate heterogeneous returns to experience. Our approach builds on the grouped fixed-effects methodology of Bonhomme and Manresa (2015), extending their framework to accommodate two-sided heterogeneity in wage growth.

While our algorithm shares the broad objective of identifying latent types in recent work (Bonhomme et al., 2019; Lentz et al., 2023; Mann, 2024), our methodological approach differs. By adopting a group fixed-effects methodology, we depart from finite mixture models, which, in practice, require researchers to specify conditional wage distributions given worker and firm types. Instead, we adopt a least-squares framework that imposes no distributional assumptions. It has two key implications. First, unlike finite mixture models that attach likelihoods to units belonging to latent types, we seek to find the optimal "hard" assignment of units to a particular type. Second, we depart from the use of the Expectation-Maximization (EM) algorithm in classifying units to types, instead using an iterative search algorithm.

3.1 Methodological Framework

Our algorithm classifies workers and firms into fixed groups based on their wage growth patterns, extending the grouped fixed-effects estimator of Bonhomme and Manresa (2015) to a setting with two-sided heterogeneity. The key innovation lies in simultaneously determining worker growth types $\rho(i)$ and firm growth types $\kappa(j)$ while estimating group-specific returns to experience. This joint classification problem presents computational challenges that we address through an iterative optimization procedure.

3.2 Objective Function

We formulate the estimation problem as the minimization of a least-squares criterion over both group assignments and parameters. Given our extended AKM specification, we seek to solve:

$$\min_{\substack{\rho:\mathcal{I}\to\{1,\ldots,K_w\},\ \kappa:\mathcal{J}\to\{1,\ldots,K_f\} \\ \beta_1,\ldots,\beta_{K_w},\ \phi_1,\ldots,\phi_{K_f} \\ \{\alpha_i\}_{i\in\mathcal{I}},\ \{\psi_{j(i,t)}\}_{j\in\mathcal{J}},\xi}} \sum_{i\in\mathcal{I}}\sum_{t\in\mathcal{T}_i}\left(\ln y_{it} - \alpha_i - \psi_{j(i,t)} - \xi' X_{it} - \beta_{\rho(i)}\text{Exp}_{it} - \sum_{j\in\mathcal{J}_i}\phi_{\kappa(j)}\text{Exp}_{it}^{j}\right)^2 \tag{7}$$

The assignment functions $\rho(\cdot) : \mathcal{I} \to \{1, ..., K_w\}$ and $\kappa(\cdot) : \mathcal{J} \to \{1, ..., K_f\}$ map each worker i and firm j to their respective growth types, with K_w and K_f denoting the number of

worker and firm groups, respectively. For ease of notation, let θ denote the set of parameters $\{\{\beta_{k_w}\}_{k_w \in K_w}, \{\phi_{k_f}\}_{k_f \in K_f}, \{\alpha_i\}_{i \in \mathcal{I}}, \{\psi_j\}_{j \in \mathcal{J}}, \xi'\}$.

This formulation embeds several important features. First, the worker assignment function $\rho(i)$ is time-invariant, reflecting our assumption that workers belong to a fixed growth type throughout their careers. Second, the firm assignment function $\kappa(j)$ similarly assigns each firm to a permanent growth type. Third, the objective function is quadratic in the parameters θ conditional on group assignments, but discontinuous in the assignment functions themselves.

The non-convex nature of this optimization problem, with its combinatorial structure over discrete assignments, is computationally challenging. The number of possible joint assignments grows exponentially with the number of workers and firms, making exhaustive search computationally infeasible even for moderately sized datasets. These challenges motivate our iterative approach, which is similar in spirit to the approach suggested by Mann (2024).

3.3 Iterative Search Algorithm

Our algorithm proceeds through coordinate descent, alternating between updating parameters and reassigning units to groups. This approach exploits the fact that, conditional on group assignments, the parameter estimation reduces to a standard OLS regression, while conditional on parameters, the optimal assignments can be determined independently for each unit.

3.3.1 Algorithm Structure

Algorithm 1. (Two-Way Clustering Iterative Search Algorithm) Starting from initial assignments $\rho^{(0)}$ and $\kappa^{(0)}$, the algorithm iterates through the following steps until convergence:

nenumerate

Parameter Update Step: Given current assignments $\rho^{(s)}$ and $\kappa^{(s)}$, estimate parameters by OLS:

$$\left(\hat{\alpha}, \hat{\psi}, \hat{\beta}, \hat{\phi}, \hat{\xi}\right)^{(s+1)} = \underset{\alpha,\psi,\beta,\phi,\xi}{\operatorname{argmin}} \sum_{i,t} \left(\ln y_{it} - \alpha_i - \psi_{j(i,t)} - \beta_{\rho^{(s)}(i)} \text{Exp}_{it} - \sum_j \phi_{\kappa^{(s)}(j)} \text{Exp}_{it}^j - \xi' X_{it} \right)^2 \tag{8}$$

Worker Assignment Step: Holding firm assignments and parameters fixed, reassign each worker to minimize the sum of squared residuals:

$$\rho^{(s+1)}(i) = \underset{k \in \{1,\ldots,K_w\}}{\operatorname{argmin}} \sum_{t \in \mathcal{T}_i} \left(w_{it}^{(s+1)} - \beta_k^{(s+1)} \text{Exp}_{it} - \sum_j \phi_{\kappa^{(s)}(j)}^{(s+1)} \text{Exp}_{it}^j \right)^2 \tag{9}$$

where $w_{it} := \ln y_{it} - \alpha_i - \psi_j - \xi X_{it}$

Firm Assignment Step: Holding worker assignments and parameters fixed, reassign each firm to minimize the objective:

$$\kappa^{(s+1)}(j) = \underset{k \in \{1,\ldots,K_f\}}{\operatorname{argmin}} \sum_{i:j \in \mathcal{J}_i} \sum_{t \in \mathcal{T}_{ij}} \left(w_{it}^{(s+1)} - \beta_{\rho^{(s+1)}(i)}^{(s+1)} \mathrm{Exp}_{it} - \phi_k \mathrm{Exp}_{it}^{j} - \sum_{j' \neq j} \phi_{\kappa^{(s)}(j')}^{(s+1)} \mathrm{Exp}_{it}^{j'} \right)^2 \quad (10)$$

where $w_{it} := \ln y_{it} - \alpha_i - \psi_j - \xi X_{it}$

The algorithm continues until the assignments stabilize or the reduction in the objective function falls below a predetermined threshold.

Initialization Strategy. The quality of the local minimum identified by our algorithm depends critically on the initial assignments. We employ a data-driven initialization strategy that leverages observable wage patterns. For worker initialization, we apply k-means clustering to the empirical distribution of individual wage trajectories. For each worker, we construct a vector containing their wages at different ages, interpolating for missing observations. We then construct the empirical cumulative distribution function of these observed wages and the average wage, grouping workers together based on these two moments. This approach groups together workers with similar life-cycle wage profiles. For firm initialization, we cluster firms based on the distribution of wage changes among job stayers. For each firm, we calculate moments of the within-firm wage growth distribution, including the mean and percentiles. The k-means algorithm then groups firms with similar wage growth characteristics.

To mitigate dependence on a single initialization, we run the algorithm from multiple starting points generated through random perturbations of the data-driven initialization as suggested by Bonhomme and Manresa (2015). The solution yielding the lowest objective function value is retained as our final estimate.

Computational Implementation. The parameter update step leverages the sparse structure of the design matrix, as each observation involves only one worker-firm pair. Modern sparse matrix algorithms enable efficient computation even with hundreds of thousands of fixed effects. The assignment steps parallelize naturally across workers, as each worker's optimal assignment can be determined independently, given the current parameters and firm assignments. One important computational constraint is that this logic does not apply to firms: a firm's optimal assignment depends, in general, on the group assignments of other firms, given current parameters and worker assignments. As a result, we must proceed sequentially, reassigning one firm at a time, updating residuals accordingly, and proceeding to the next firm. We start from the

largest firm and proceed to the smallest firm in each iteration.

3.4 Relationship to Existing Methods

Our approach synthesizes insights from the clustering and panel econometrics literature while maintaining crucial distinctions from related methodologies. The connection to Bonhomme and Manresa (2015) is direct: in the special case where either $K_w = N$ (each worker is their own type) or $K_f = J$ (each firm is their own type), our estimator collapses to their grouped fixed-effects estimator applied to one dimension of heterogeneity.

The extension to two-sided classification, however, introduces novel challenges. Unlike the one-sided case where the k-means algorithm provides a natural solution, the joint classification problem lacks a standard algorithmic approach. Our coordinate descent strategy, alternating between worker and firm assignments, provides a computationally tractable solution that scales to large administrative datasets.

Our approach differs from finite mixture methods such as Bonhomme et al. (2019), Lentz et al. (2023), and Mann (2024) in several respects. We estimate latent groups through hard-assignment least squares, minimizing the within-group sum of squared residuals via alternating updates and assigning each unit deterministically to one group. This objective does not require a fully specified parametric error distribution. In contrast, finite mixture approaches maximize a likelihood with soft (probabilistic) memberships, typically under parametric component densities, and are commonly estimated by EM.

3.5 Selection of Group Numbers

The choice of K_w and K_f involves a trade-off between model flexibility and estimation precision. Increasing the number of groups allows for richer heterogeneity patterns but reduces the number of observations per group, potentially compromising the precision of group-specific parameter estimates.

For firm groups, we employ a data-driven selection procedure based on the Calinski-Harabasz index, which measures the ratio of between-group to within-group variation, similarly to Lentz et al. (2023). We estimate models with different values of K_f and select the specification maximizing this index, indicating well-separated, internally homogeneous groups. For worker groups, computational constraints impose a practical upper bound. Given the scale of our data and the computational complexity of the algorithm, we set $K_w = 20$, which provides sufficient flexibility to capture the heterogeneity in worker growth patterns while maintaining computational feasibility.

To ensure our results are not driven by our particular choice of group numbers, we conduct robustness checks by keeping the number of worker groups fixed while varying the number of firm groups between 20 and 100, and vice versa. These exercises test whether our findings are sensitive to the dimensional specifications on either side. We present these robustness results in Section 4.

4 Algorithm Validation and Performance

This section validates our two-way clustering algorithm through Monte Carlo simulations under a particular data-generating process (DGP). We examine four scenarios that vary along two dimensions: whether workers and firms match assortatively or randomly, and whether wage growth heterogeneity is driven primarily by workers or firms. These simulations demonstrate its ability to recover both the classification of units into types and the resulting variance decompositions. Finally, we assess the sensitivity of our results with respect to the choice of cluster numbers.

4.1 Monte Carlo Experiments

To assess the performance of our two-way clustering algorithm, we conduct Monte Carlo simulations that validate it as a practical approximation to the two-way individual fixed effects framework.[9] Our simulation approach serves two primary objectives: first, to evaluate the algorithm's ability to correctly classify workers and firms into discrete growth types when the true DGP features continuous individual-level heterogeneity; and second, to assess whether the resulting variance decompositions accurately capture the true sources of life-cycle wage inequality.

We generate wage data from the following two-way individual fixed effect model based on equation (4):

$$\ln y_{i,t} = \alpha_i + \psi_{j(i,t)} + \beta_i \operatorname{Exp}_{i,t} + \sum_j \phi_j \operatorname{Exp}_{i,t}^j + \epsilon_{i,t} \tag{11}$$

where, in addition to worker and firm fixed effects (α_i and $\psi_{j(i,t)}$), each worker i receives an individual-specific return to one year of experience β_i and at each firm j, workers receive a firm-specific return to one year of tenure ϕ_j. Both β_i and ϕ_j are drawn from continuous distributions. We then apply our clustering algorithm to jointly classify workers and firms into discrete growth types and estimate the group-specific returns to experience $\beta_{\rho(i)}$ and $\phi_{\kappa(j)}$. This design allows us

[9] While the individual FE approach is feasible in principle, it is computationally costly and produces noisy estimates, making it unclear what can be learned from such estimates in practice.

to evaluate whether the estimated discrete groupings successfully approximate the underlying continuous distributions and whether estimated variance decompositions accurately reflect the true variance components over the life cycle.

Table 1 summarizes our simulation design. Panel A specifies parameters common to all scenarios, including the distributions of worker and firm fixed effects and idiosyncratic shocks. Panel B outlines four scenarios that vary along two key dimensions: whether workers and firms match assortatively or randomly, and the relative importance of worker versus firm contributions to wage growth. Each simulated dataset contains 500,000 individuals working in 5,000 firms observed over 40 periods.

Table 1: Monte Carlo Simulation

Panel A: Common Parameters Across All Scenarios	
Parameter	Value
Number of Workers ($\|\mathcal{I}\|$)	$500,000$
Number of Firms ($\|\mathcal{J}\|$)	$5,000$
Time periods ($\|\mathcal{T}\|$)	40
Worker growth types ($\beta_{\rho(i)}$)	$\beta_i \sim N(0.08,\ 0.03)$
Firm growth types ($\phi_{\kappa(j)}$)	$\phi_j \sim N(0.03,\ 0.015)$
Worker fixed effects (α_i)	$\alpha_i \sim N(0,\ 0.6)$
Firm fixed effects (ψ_j)	$\psi_j \sim N(0,\ 0.25)$
Idiosyncratic shocks (ε_{it})	$\varepsilon_{it} \sim N(0,\ 0.15)$
Panel B: Scenario-Specific Features	
Scenario	Description
1: No Sorting	Random matching; job switching prob. $= 0.10$; baseline (β_i, ϕ_j).
2: Sorting	Assortative matching with $\lambda = 0.5$; intensity grows at rate 0.07 per period (max 2.0); baseline (β_i, ϕ_j).
3: Worker-Dominated	Random matching; job switching prob. $= 0.10$; $\beta_i \sim N(0.08,\ 0.05)$, $\phi_j \sim N(0.03,\ 0.005)$.
4: Firm-Dominated	Random matching; job switching prob. $= 0.10$; $\beta_i \sim N(0.08,\ 0.005)$, $\phi_j \sim N(0.03,\ 0.025)$.

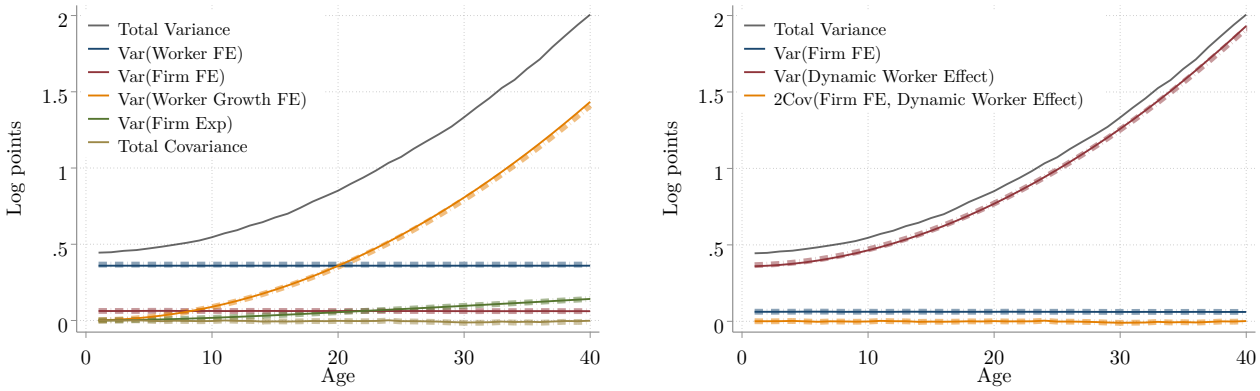
Notes: This table shows parameters in Monte Carlo simulations. Panel A lists parameters common to all simulations. Panel B describes the four scenarios. In Scenario 2, sorting intensity increases at rate 0.07 per period up to a maximum multiplier of 2.0. Scenarios 3 and 4 modify the standard deviations of growth heterogeneity while keeping means unchanged.

Scenario 1 serves as our baseline with random worker-firm matching. Scenario 2 introduces dynamic assortative sorting that strengthens over the life cycle. Scenarios 3 and 4 examine cases where wage growth heterogeneity is dominated by either workers or firms, respectively, while maintaining random matching. We now discuss each scenario in detail.

4.1.1 Scenario 1: No Sorting

Scenario 1 establishes our baseline with random worker-firm matching. Workers are initially assigned to firms uniformly at random. In each subsequent period, job-to-job transitions occur randomly with probability 0.10, ensuring no systematic relationship between worker characteristics (α_i, β_i) and firm characteristics (ψ_j, ϕ_j).

Figure 2: Variance Decomposition over the Life Cycle: No Sorting



(a) All Components

(b) Dynamic AKM

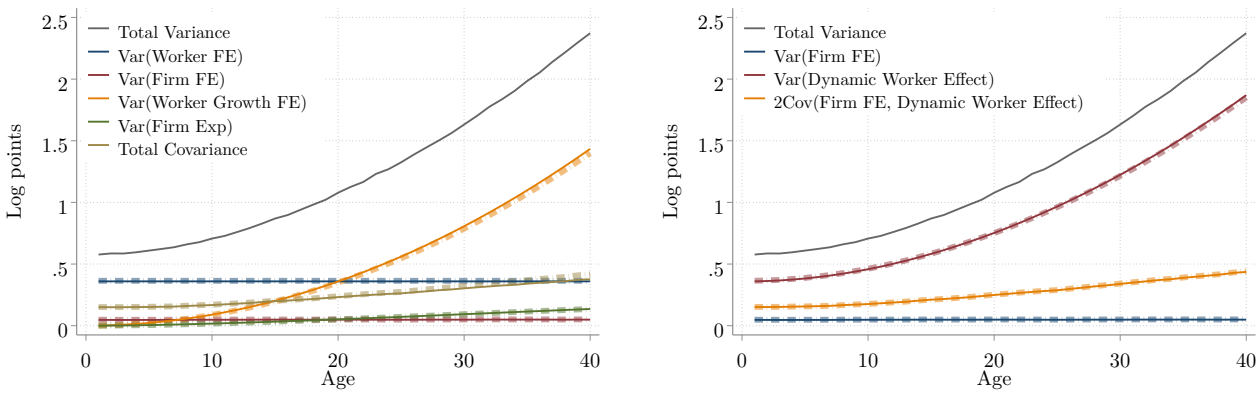
Notes: Both panels decompose log wage variance over the life cycle in the no-sorting simulation. The black line in each panel represents total variance. Panel (a) shows the full decomposition: Panel (b) aggregates all worker-side components into a single dynamic worker effect.

4.1.2 Sorting

We consider two DGPs that differ in their treatment of worker-firm matching.

The first DGP incorporates dynamic sorting, where workers progressively sort to firms based on their characteristics, with matching strength increasing over the life cycle. This captures the empirical observation that high-ability workers increasingly concentrate at high-paying firms as they gain experience.

Figure 3: Variance Decomposition over the Life Cycle: Sorting



(a) All Components

(b) Dynamic AKM

Notes: Both panels decompose log wage variance over the life cycle in the sorting simulation. The black line in each panel represents total variance. Panel (a) shows the full decomposition. Panel (b) aggregates all worker-side components into a single dynamic worker effect.

Both DGPs feature 400 workers and 80 firms observed over 40 periods, with heterogeneous worker growth components ($\beta_i \sim N(0.08, 0.03)$) and firm growth components ($\phi_j \sim N(0.03, 0.015)$). Worker fixed effects are drawn from $N(0, 0.6)$ and firm fixed effects from $N(0, 0.25)$, with idiosyncratic shocks following $N(0, 0.15)$.

The sorting DGP implements assortative matching that strengthens over time. Initial matching follows a probability function that depends on the distance between worker and firm types, with matching strength governed by parameter $\lambda = 0.5$. As workers accumulate experience, the sorting intensity increases at rate 0.07 per period up to a maximum multiplier of 2.0. In contrast, the no-sorting DGP assigns workers to firms uniformly at random, conditional on job switching.

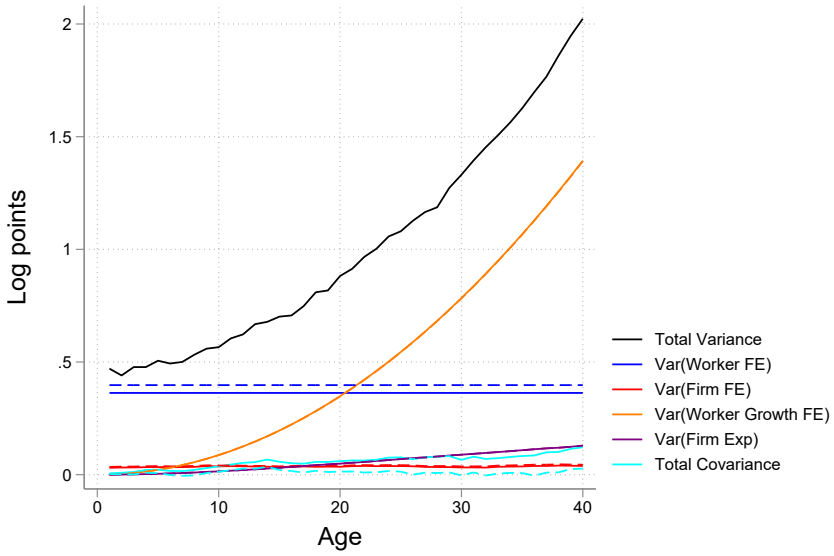
Our algorithm performs well in recovering the underlying heterogeneity structure. In the sorting DGP, the estimated worker groups explain 85.1% of the true variance in worker growth components (β_i), while firm groups explain 52.1% of the variance in firm growth components (ϕ_j). Performance improves further in the no-sorting DGP, with 93.2% and 62.9% of variance explained for workers and firms, respectively. The higher explanatory power in the no-sorting case reflects the reduced confounding between sorting patterns and unit heterogeneity.

Table 2: Algorithm Classification Performance Across All Scenarios

Panel A: No Sorting	
Variance in β_i explained by worker groups	89.9%
Variance in ϕ_j explained by firm groups	62.9%
Panel B: Sorting	
Variance in β_i explained by worker groups	85.1%
Variance in ϕ_j explained by firm groups	52.1%
Panel C: Worker-Dominated Learning	
Variance in β_i explained by worker groups	*[to be computed]*
Variance in ϕ_j explained by firm groups	*[to be computed]*
Panel D: Firm-Dominated Learning	
Variance in β_i explained by worker groups	*[to be computed]*
Variance in ϕ_j explained by firm groups	*[to be computed]*

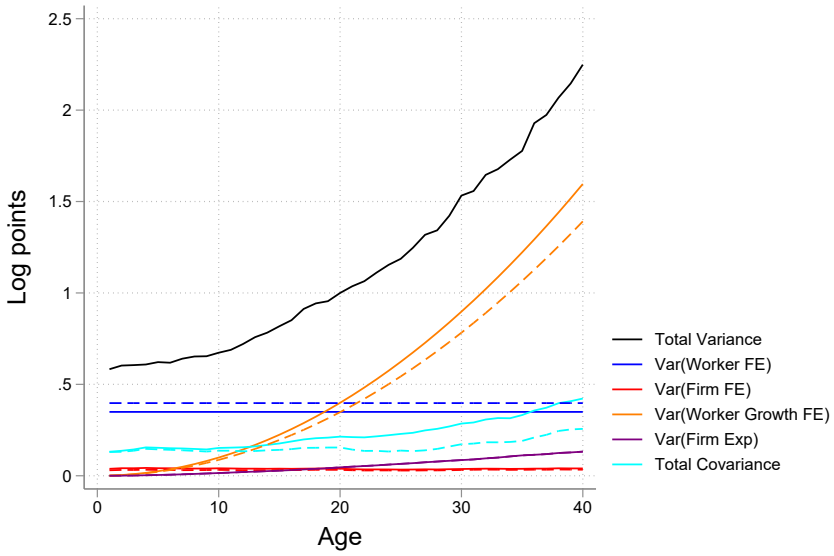
Notes: This table reports the proportion of true variance in growth components explained by estimated group classifications across all four scenarios. Worker groups capture variance in individual-specific growth rates (β_i), while firm groups capture variance in firm-specific growth rates (ϕ_j). Higher percentages indicate better recovery of the true heterogeneity structure.

Figure 4: Variance decomposition over the life-cycle: variance components



Notes: This figure shows the decomposition of log wage variance into its constituent components over the life cycle in the no-sorting simulation. The black line represents total variance. Solid colored lines show estimated variance components from our algorithm: $Var(\hat{\alpha}_i|a)$ in blue, $Var(\hat{\psi}_j|a)$ in red, $Var(\hat{\beta}_{\hat{\rho}(i)}Exp_{i,t}|a)$ in orange, $Var(\sum_j \hat{\phi}_{\hat{\kappa}(j)}Exp^j_{i,t}|a)$ in purple, and the sum of all covariance components in cyan. Dashed lines of the same colors represent the true variance components from the data-generating process. The close alignment between solid and dashed lines demonstrates the algorithm's ability to accurately recover the variance decomposition even without observing true types.

Figure 5: Variance decomposition over the life-cycle: variance components



Notes: This figure presents the variance decomposition under the sorting DGP. The black line shows total variance, which increases more steeply than in the no-sorting case. Solid colored lines represent estimated components: Var(Worker FE) in blue remains constant, Var(Firm FE) in red shows slight variation, Var(Worker Growth FE) in orange increases substantially with age, Var(Firm Exp) in purple contributes modestly, and Total Covariance in cyan increases notably. Dashed lines show true values. The growing importance of covariance terms reflects the strengthening sorting between high-growth workers and high-wage firms over the life cycle.

4.1.3 Scenario 3: Worker-Dominated Learning

Simulate under scenarios where worker vs. firm effects dominate (e.g., only worker heterogeneity in learning; only firm heterogeneity in learning).

- Show the algorithm recovers the correct "order" (relative importance of worker vs. firm growth).
- Use simulated data from *our statistical model* (to avoid misspecification bias), not from Victoria Gregory's economic model for the paper.

4.1.4 Scenario 4: Firm-Dominated Learning

4.2 Robustness to Cluster Number Specification

- Demonstrate stability across cluster pairs (e.g., 20/50, 50/20, 20/30, 30/20).
- Working hypothesis: 20 groups on both sides should be "enough."

5 Data

This section describes our data and sample construction. We first introduce the Veneto Worker History dataset, detailing our sample selection procedures. We then present descriptive statistics characterizing workers, firms, and wage dynamics in the Veneto labor market.

5.1 Data

Data. We use social security administrative data, the Veneto Worker History (VWH) dataset, which contains the entire working population and private firms in the region of Veneto, Northeast Italy, from 1975 to 2001. Veneto is the fourth most populous region in Italy, with a population of about five million in 2012. The dataset includes unique worker and firm identifiers, employment histories, workers' demographic information, and firm-level variables.[10] The VWH contains detailed worker-level information, including annual earnings (without top-coding), the number of days worked that year for every job spell, contract type, occupation (categorized as white-collar, blue-collar, manager, or apprentice), gender, age, birthplace, and seniority within the firm. For firms, the data includes national tax identifiers, location, start and closure dates, and industry classification.

Variable construction and sample selection. We focus on data from the years 1984-2001 as prior information tends to be lower quality (Kline et al., 2020). To construct the person-year panel, we follow standard sample selection procedures in the AKM literature. First, we restrict our sample to workers aged 25 or older. In particular, we consider workers born in and after 1959 in the Veneto region, allowing us to observe their labor market outcomes from age 25 through mid-forties. We drop workers with missing gender and employment spells with no firm identifiers. We also exclude part-time workers (identified from contract classifications), apprentices, and public sector workers because their wages cannot be compared to regular full-time employment. To generate unique worker-firm assignments in each year, we keep only a worker's primary job if he or she holds multiple jobs in a year. We define a worker's primary job as the job with the highest annual earnings.[11] Annual earnings are adjusted for inflation using 2003 as the base year. Following Kline et al. (2020), we drop workers with wage change observations that yield changes in excess of 1 or -1. For each worker, we observe both days worked and total earnings (which include overtime payments) from their primary job. We construct average daily wages by dividing total earnings by the number of days worked. Finally, due to the identification requirement in the AKM framework, we restrict the sample to the largest connected set (Abowd

[10] Previous papers that have used the VWH data include Card et al. (2014), Battisti (2017), Bartolucci et al. (2018), Serafinelli (2019), Kline et al. (2020), and Arellano-Bover and Saltiel (2024).

[11] We break ties for very few cases where workers hold multiple jobs with the same highest annual earnings by randomly choosing the primary job.

et al., 1999).

5.2 Descriptive Statistics

Table 3 presents descriptive statistics for the sample of workers and firms in Veneto used in our analysis. We have nearly 7 million person-year observations, covering 1.1 million workers and 400 thousand firms. The sample is 65% male. On average, workers are about 30 years old. The cohorts we observe from age 25 through their mid-40s spend on average 5.2 full-year equivalents employed in the formal sector, and spend on average 3.5 years at each firm they are employed at. The majority of workers are employed in blue-collar occupations (65 percent), with the remainder mainly working in white-collar positions (34 percent). Workers in our sample receive 130.56 euros (in 2003 prices) per day in their primary jobs, and their primary jobs on average last for over 60 percent of a year.

Table 3: Summary Statistics

	(1) Mean	(2) SD
Worker Characteristics		
Fraction of male	0.65	0.48
Worker's age	30.29	4.18
Tenure since age 25	3.48	3.00
Experience since age 25	5.27	3.85
Fraction of blue-collar	0.65	0.48
Fraction of white-collar	0.34	0.47
Fraction of manager	0.01	0.09
Job, Wage Characteristics		
Primary job daily wages	130.56	164.60
Job duration (in days)	237.26	104.19
Annual wage growth	0.02	0.12
Within firm wage growth	0.02	0.11
Between firm wage growth	0.01	0.21
Firm Characteristics		
Firm size	11.68	80.14
Number of person-year observations:		6,910,126
Number of person-year observations in largest connected set:		5,362,684
Number of unique workers:		1,136,248
Number of unique firms:		397,603
Number of unique firms in largest connected set:		114,985
Number of 'moves' in largest connected set:		438,068

Notes: The table reports means and standard deviations of each variable based on the Veneto Worker History Panel. Sample restrictions are described in Section 5.

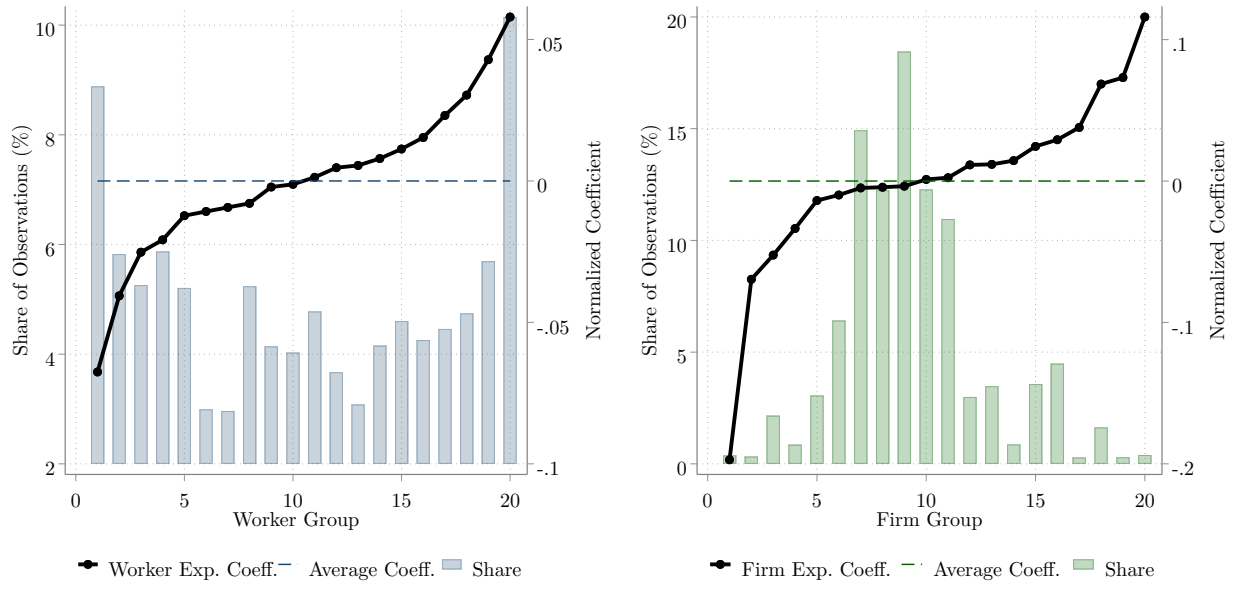
Table 3 also shows that workers in our sample experience an average annual wage increase of 0.02 log points in Veneto. Wage growth is meaningful both within and between jobs, as average within- and between-firm annual wage increases reach 0.02 and 0.01 log points, respectively. As expected, firms are small in the labor market of Veneto, with a mean of 12 employees and a standard deviation of 80.

6 Two-sided Heterogeneity in Returns to Experience

This section presents our baseline estimates of heterogeneous returns to experience across worker and firm groups and validates our clustering algorithm. We first present estimates showing substantial dispersion in returns to experience, with important differences in how workers and firms are distributed across growth types. We then validate our two-way clustering algorithm by comparing our group-level estimates with individual-level fixed effects obtained from an alternative estimation procedure.

6.1 Main Estimates

Figure 6: Returns to experience by worker groups and firm groups



Notes: This figure shows the returns to experience by worker groups and firm groups in Veneto. The black dotted lines plot the estimated returns to experience coefficients for each group (right y-axis), normalized such that the average equals zero. The bars show the share of person-year observations in each group (left y-axis). Both worker and firm groups are ranked by the magnitudes of their respective returns to experience coefficients.

Figure 6 displays estimates of equation (5), presenting our baseline results on returns to experiences acquired by different worker groups and at different firm groups. The black dotted line plots the estimated returns to experience coefficient for each group (right y-axis), normalized such that the average coefficient equals zero (shown by the horizontal dashed line).[12] The bars show the share of person-year observations assigned to each group (left y-axis)

Our main finding is that substantial heterogeneity in returns to experience exists across both worker groups and firm groups. Worker groups exhibit meaningful dispersion in returns to experience. The coefficients range from approximately -0.06 (6 percentage points below average) for group 1 to 0.05 (5 percentage points above average) for group 20, yielding a spread of 11 percentage points between the lowest and highest groups.[13] Importantly, workers are relatively balanced across groups, with each group containing 3-10% of person-year observations, and notably, the extreme groups (groups 1 and 20) contain relatively larger shares of workers.

Firm groups display an even wider range in estimated coefficients, spanning from approximately -0.2 for group 1 to 0.1 for group 20 — a spread of 30 percentage points. However, this wider range does not translate into correspondingly large wage growth coefficient dispersion in practice: the vast majority of person-year observations are concentrated at firms in the middle groups (groups 7-10) with coefficients close to zero, with some groups employing up to 18% of all person-years. In contrast, the extreme firm groups with very high or very low returns to experience collectively account for a much smaller share of employment.

These distributional patterns have important implications for wage inequality. Despite the narrower coefficient range for worker groups, the observation-weighted dispersion — the variation in worker experience coefficients —is substantial because workers are spread relatively evenly across high-growth and low-growth types, with notable concentration at the extremes. In contrast, despite the wider coefficient range for firm groups, the observation-weighted dispersion is more limited because most workers are employed at firms offering near-average returns to experience. This suggests that worker heterogeneity plays a more important role than firm heterogeneity in determining life-cycle wage inequality.

[12] We normalize the coefficients by subtracting the average return to experience from each group's coefficient. This normalization facilitates comparison across groups while preserving the dispersion in returns to experience.

[13] These normalized coefficients represent deviations from the average return to experience. Negative coefficients indicate that, compared to workers with identical gender and age, this group of workers accumulates lower returns to experience than the average in that demographic group.

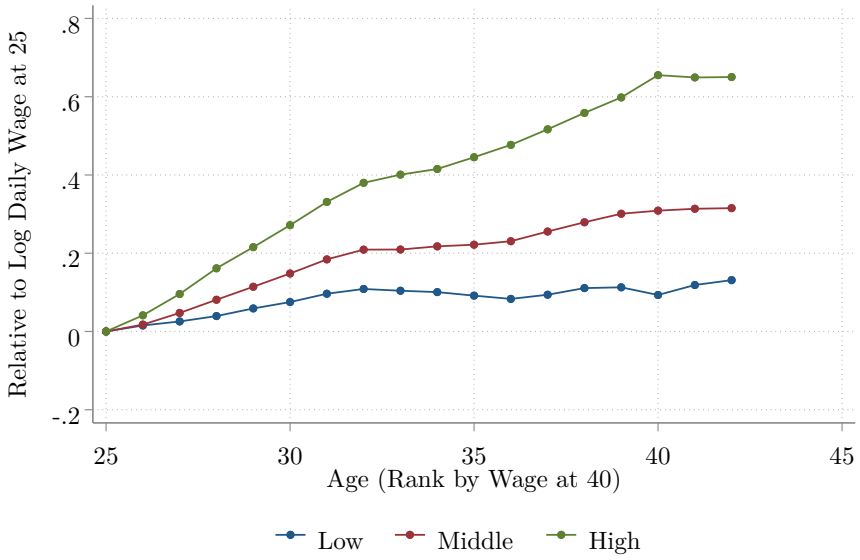
7 Wage Inequality Over the Life Cycle

7.1 Decomposing Life-Cycle Wage Inequality

Equation (5) provides a simple framework to measure how four key components — worker fixed effect, firm fixed effect, worker growth component, and firm experience component — contribute to life-cycle wage inequality. Using our framework, we perform two life-cycle wage decompositions: one focusing on wage levels and another on wage growth.

To implement the decomposition, we first construct a balanced panel of male workers from a single cohort, which allows us to track the same group of workers over their careers and minimize changes in worker composition. Second, we rank workers into three distinct terciles $i \in \{g_{min}, g_{med}, g_{max}\}$ based on their wage at age 40.[14] Within each tercile, we then calculate the mean estimates for each wage component. The differences in these means between terciles reveal how each component contributes to overall wage inequality. Figure 7 displays the mean life-cycle wage profiles for each tercile, adjusted for year fixed effects.

Figure 7: Life-Cycle Wage Growth



Notes: This figure shows the life-cycle wage growth profiles (adjusted for year fixed effects) for workers in three terciles in Veneto. It is calculated by taking the differences between the log daily wage at later ages and the log daily wage at the age of 25. Workers are classified into three groups based on their wage rank at age 40: low (33rd percentile), middle (66th percentile), and high (99th percentile). Log daily wages are computed using primary job annual earnings divided by days worked, deflated by the Italian 2003 price level.

[14] We choose wage at age 40 as our reference point since previous literature documents that the wage-age profile is basically flat at age 40, which is approximately true in our context, Veneto region.

Life-cycle Wage Level Decomposition. We denote the average log daily wage for tercile g at age a as $\mathbb{E}a, g \ln y_{i,t}$. We decompose the wage level differences between each tercile g and the lowest tercile g_{min} as follows:

$$\begin{aligned}\mathbb{E}_{a,g} \ln y_{i,t} - \mathbb{E}_{a,g_{\min}} \ln y_{i,t} = & \underbrace{(\mathbb{E}_{a,g}\hat{\alpha}_i - \mathbb{E}_{a,g_{\min}}\hat{\alpha}_i)}_{\text{Diff. in Composition Effect}} + \underbrace{(\mathbb{E}_{a,g}\hat{\psi}_j - \mathbb{E}_{a,g_{\min}}\hat{\psi}_j)}_{\text{Diff. in Firm Rents}} \\ & + \underbrace{(\mathbb{E}_{a,g}\hat{\beta}_i \operatorname{Exp}_{i,t} - \mathbb{E}_{a,g_{\min}}\hat{\beta}_i \operatorname{Exp}_{i,t})}_{\text{Diff. in Returns to Worker Exp. Component}} \\ & + \underbrace{(\mathbb{E}_{a,g} \sum_{\kappa(j)} \hat{\phi}_{\kappa(j)} \operatorname{Exp}_{i,t}^{\kappa(j)} - \mathbb{E}_{a,g_{\min}} \sum_{\kappa(j)} \hat{\phi}_{\kappa(j)} \operatorname{Exp}_{i,t}^{\kappa(j)})}_{\text{Diff. in Returns from Firm Exp. Component}} \\ & + \mathbb{E}_{a,g}\text{Residual}\end{aligned} \tag{12}$$

Equation (12) decomposes wage differences between tercile g and the lowest tercile into four components. The first term compares worker fixed effects α_i, which capture unobserved attributes at labor market entry — such as latent ability or skills — that are rewarded equally across firms. The second term compares the firm fixed effects ψ_j across the jobs held by each tercile, revealing whether higher-wage workers sort into firms paying higher premiums. The third and fourth terms measure differences in accumulated returns to experience: the worker growth component $\beta_i \text{Exp}_{i,t}$ captures portable returns to experience that workers carry across employers, while the firm growth component $\sum_{\kappa(j)} \phi_{\kappa(j)} \text{Exp}_{i,t}^{\kappa(j)}$ reflects returns specific to experience acquired at particular firm types. Positive values indicate that higher tercile workers possess greater latent ability at entry, work at premium-paying firms, or have accumulated steeper returns to experience through either channel. The residual captures unexplained wage differences.

Life-cycle Wage Growth Decomposition. To decompose the life-cycle wage growth for each tercile relative to age 25, we replace the expectation $\mathbb{E}$ with the difference Δ in Equation (12), defining the difference in average log wages at age a relative to age 25 for each tercile g as $\Delta_{a,g} \ln y_{i,t} := \mathbb{E}[\ln y_{i,t} | age = a, group = g] - \mathbb{E}[\ln y_{i,t} | age = 25, group = g]$. The resulting decomposition (Equation (13)) quantifies the contributions of worker composition, firm wage premia, worker growth component, firm growth component, and the residual component to overall life-cycle wage growth relative to age 25.

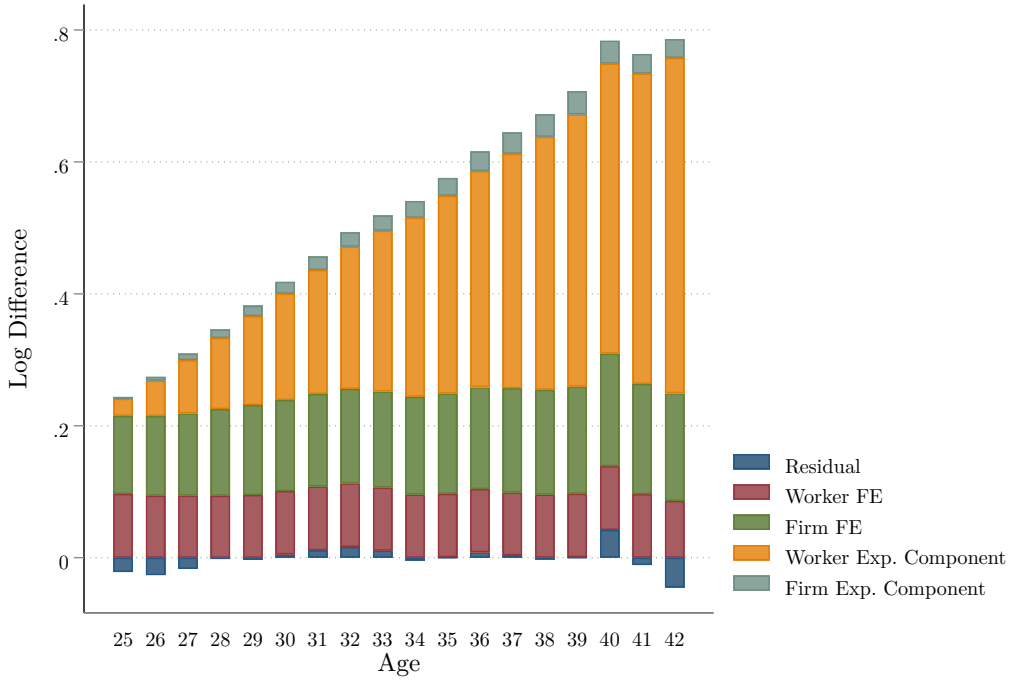
$$\Delta_{a,g} \ln y_{i,t} - \Delta_{a,g_{min}} \ln y_{i,t} = \underbrace{(\Delta_{a,g}\alpha_i - \Delta_{a,g_min}\alpha_i)}_{\text{Diff in Worker FE}} + \underbrace{(\Delta_{a,g}\psi_j - \Delta_{a,g_min}\psi_j)}_{\text{Diff in Firm FE}} + \underbrace{\Delta_{a,g}\beta_i Exp_{i,t} - \Delta_{a,g_min}\beta_i Exp_{i,t}}_{\text{Diff in Returns to Worker Exp. Component}} + \underbrace{\Delta_{a,g}\sum_{\kappa(j)} \phi_{\kappa(j)} Exp_{i,t}^{\kappa(j)} - \Delta_{a,g_min}\sum_{\kappa(j)} \phi_{\kappa(j)} Exp_{i,t}^{\kappa(j)}}_{\text{Diff in Returns to Firm Exp. Component}} \tag{13}$$

Figure 8 presents the decomposition results of wage differences across the life-cycle between the highest and lowest terciles of workers. Panel (a) shows the life-cycle wage level decomposition at each age, while panel (b) decomposes life-cycle wage growth relative to age 25. In Panel (a), there is a widening gap in wages over the life-cycle between these two terciles of workers. This growing inequality is primarily driven by increasing differences in the average worker growth component (orange bar) across two groups, suggesting that workers in the top tercile experience much steeper wage profiles as they accumulate experience in the labor market compared to the bottom tercile.

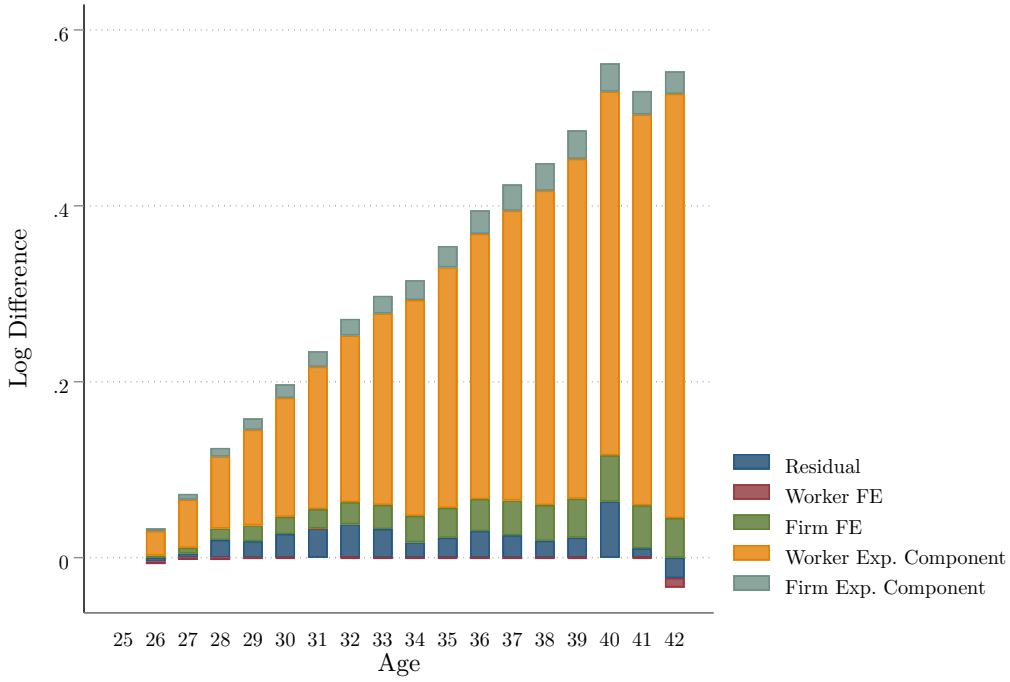
Panel (b) takes the difference in each component at later ages relative to the respective levels at age 25, allowing us to decompose the growth in the wage gap over the life cycle into the contributions from each component. By subtracting the age 25 level, panel (b) removes the initial differences that already exist between terciles at the age of 25. What remains is the differential growth in each component over time. Here, the role of the worker growth component is even more stark. The worker growth component accounts for the vast majority of the divergence in wage trajectories between the top and bottom terciles. Intuitively, this means that top-tercile workers see much steeper returns for each additional year of experience compared to bottom-tercile workers, leading to a widening wage growth gap.

The other two components — worker fixed effects and firm fixed effects — are relatively stable over time, as suggested in Panel (a). Both components play a minimal role in the life-cycle wage decomposition. The worker fixed effect (red bar) remains fairly constant, reflecting the fact that the analysis is based on a balanced panel of workers, so worker composition does not change. Moreover, the nature of the worker fixed effect as a time-invariant component means it is not expected to evolve as workers age. The modest increase in the firm effects (green bar) and firm growth component (teal bar) in Panel (b) suggests that top-tercile workers are more likely to sort into firms with higher firm premia and work at firms offering steeper returns to experience over time, contributing to the differences in wage growth patterns over the course of their careers.

Figure 8: Life-Cycle Wage Decomposition



(a) Wage Level Decomposition, Top and Bottom Terciles



(b) Wage Growth Decomposition, Top and Bottom Terciles

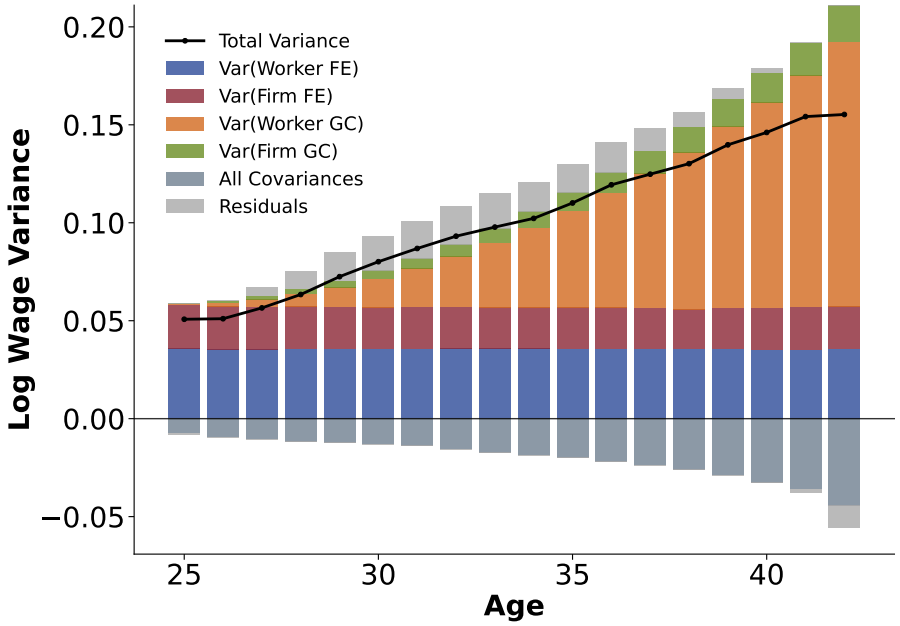
Notes: These figures show the decomposition of life-cycle wages and wage growth between workers in the top and bottom terciles. Panel (a) shows the life-cycle wage level decomposition at each age, while panel (b) decomposes life-cycle wage growth relative to age 25.

7.2 Variance Decomposition

We next quantify how much of the variance of wages is accounted for by heterogeneous worker returns to experience and heterogeneous firm returns to experience. We build on AKM literature (Card et al., 2013; Alvarez et al., 2018) and perform multiple variance decomposition exercises to illuminate the sources of life-cycle wage inequality.

The evolution of wage inequality over the life cycle represents one of the most striking patterns in labor economics. Workers who begin their careers with similar wages experience dramatically divergent trajectories, with some climbing steadily while others stagnate. Our decomposition framework allows us to precisely quantify the sources of this divergence, separating the contributions of worker heterogeneity, firm heterogeneity, and their dynamic interactions over workers' careers.

Figure 9: Variance Decomposition



Notes: This figure decomposes the variance of log wages over the life cycle. The black line represents total variance. The stacked bars show $\text{Var}(\hat{\alpha}_i)$ (Worker FE) in blue, $\text{Var}(\hat{\psi}_j)$ (Firm FE) in red, $\text{Var}(\hat{\beta}_{\rho(i)}\,\text{Exp}_{i,t})$ (Worker Growth Component) in orange, $\text{Var}(\sum_{\kappa(j)}\hat{\phi}_{\kappa(j)}\,\text{Exp}_{i,t}^{\kappa(j)})$ (Firm Growth Component) in green. All covariance terms are combined in dark gray, and residuals in light gray.

Figure 9 reveals a critical finding: heterogeneity in worker growth effects is the main driver of the increase in wage inequality over the life cycle. The variance decomposition demonstrates that while worker and firm fixed effects remain relatively stable across ages, reflecting the time-invariant nature of these components, the variance attributable to the worker growth component expands dramatically. The worker growth component, which captures heterogeneous returns to experience across workers irrespective of their employer, accounts for the majority of the

tripling in wage variance observed between ages 25 and 42.

The dominance of worker growth heterogeneity becomes even more striking when we consider its scale relative to other components. By age 42, the variance of worker growth effects ($\hat{\beta}_{\rho(i)} \text{Exp}_{i,t}$) reaches 0.10 log points, dwarfing the contribution of firm growth effects, which peaks at merely 0.02 log points. This substantial dispersion in the worker growth component—as shown in earlier analysis documenting worker experience effects, has a larger dispersion, translating directly into widening wage gaps as workers accumulate experience over their careers.

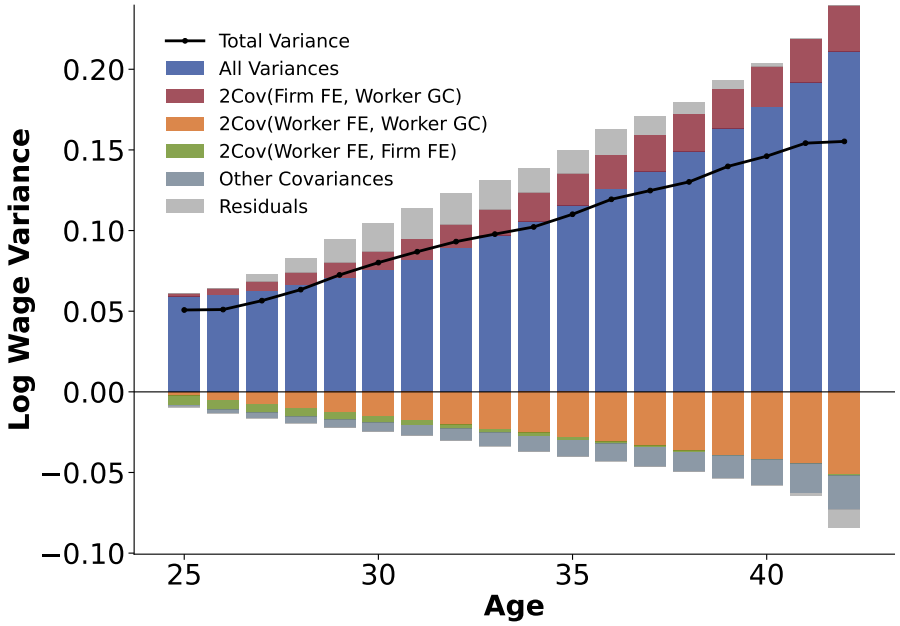
In contrast, the comparatively modest contribution of firm growth effects to rising inequality reflects two distinct features of the Veneto labor market. First, the dispersion in firm experience effects is inherently smaller than that of worker experience effects, suggesting relatively limited heterogeneity in how different firms reward tenure and experience. Second, and perhaps more importantly, worker mobility patterns result in substantial compression of firm-specific growth effects: workers are not stranded at firms of particular growth types for the entirety of their careers. The typical worker in our sample works for multiple employers over the observation period, effectively averaging out exposure to different firm growth environments and limiting the cumulative impact of any single firm's growth trajectory on life-cycle wage dispersion.

Figure 10 presents two key facts about the covariance structure underlying life-cycle wage inequality. First, and most strikingly, we document that "high growth" workers become increasingly likely to work at "high wage" firms as they progress through their careers. The covariance between worker growth effects and firm fixed effects (red bar) starts near zero at age 25 but grows substantially over the life cycle, contributing positively to overall wage dispersion. This pattern reveals a dynamic sorting process: workers with steeper experience profiles progressively match with firms offering higher firm wage premia, amplifying inequality through both channels.

The mechanism underlying this sorting pattern potentially reflects the differential ability of high-growth workers to climb the job ladder. For instance, workers who experience rapid human capital accumulation or possess strong bargaining abilities may become increasingly attractive to high-paying employers. Through on-the-job search, these workers may gradually sort into firms with more generous compensation policies, creating a compounding advantage that extends beyond their individual growth trajectories.

The second key fact emerging from Figure 10 is the negative and increasingly pronounced covariance between worker fixed effects and worker growth effects. This negative association (orange bar) indicates that workers who command high initial wages (high $\hat{\alpha}_i$) tend to have flatter wage trajectories (low $\hat{\beta}_{\rho(i)}$), partially offsetting what would otherwise be even more dramatic increases in inequality. This suggests that the labor market exhibits some mean-

Figure 10: Covariance Components



Notes: This figure decomposes the total variance of log wages into its covariance components over the life cycle. The black line shows total variance, while the blue bar represents the sum of all variance components (worker FE, firm FE, worker growth, and firm growth variances combined). Covariance terms are shown as $2\,\mathrm{Cov}(\hat{\psi}_j, \hat{\beta}_{\rho(i)}\,\mathrm{Exp}_{i,t})$ (Firm FE with Worker Growth) in red, $2\,\mathrm{Cov}(\hat{\alpha}_i, \hat{\beta}_{\rho(i)}\,\mathrm{Exp}_{i,t})$ (Worker FE with Worker Growth) in orange, $2\,\mathrm{Cov}(\hat{\alpha}_i, \hat{\psi}_j)$ (Worker FE with Firm FE) in green, and other minor covariances in dark gray, and residuals in light gray.

reverting tendencies: workers who begin with advantages in terms of initial productivity or bargaining position experience more modest growth, while those starting from lower initial wages can partially catch up through steeper growth trajectories.

The negative correlation between worker-level effects and worker growth effects may reflect several underlying economic forces. High initial wage workers may have already revealed their productivity early in their careers, leaving less scope for learning and wage growth. Alternatively, firms may structure compensation to provide steeper growth incentives for initially lower-paid workers while offering flatter but higher profiles to those with demonstrated early-career productivity. Regardless of the precise mechanism, this negative covariance serves as a moderating force, preventing life-cycle inequality from expanding even more dramatically than observed.

7.3 Dynamic AKM Decomposition

To provide a complementary perspective on the sources of life-cycle wage inequality, we now turn to a dynamic extension of the canonical AKM decomposition. The key innovation of this exercise is to define a time-varying worker effect that captures all portable components of wages, those elements that workers carry with them as they move between employers. Specifically, we define the dynamic worker effect as:

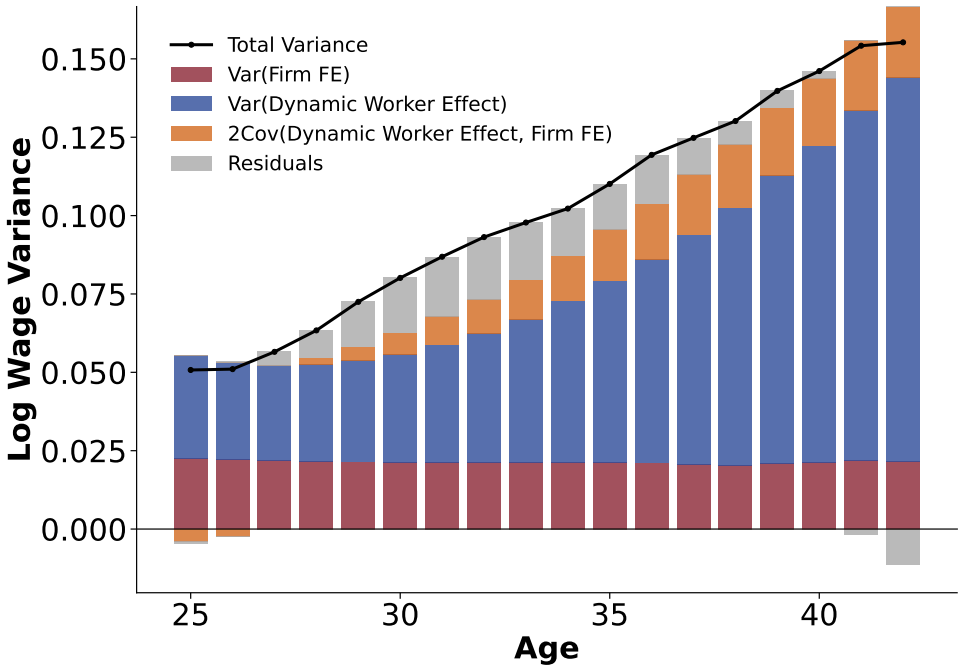
$$\alpha_{i,t} = \hat{\alpha}_i + \hat{\beta}_{\rho(i)} \, \text{Exp}_{i,t} + \sum_j \hat{\phi}_{\kappa(j)} \, \text{Exp}_{i,t}^j \tag{14}$$

This formulation recognizes that, through the lens of our model, all three components, the worker fixed effect, the worker-specific returns to experience, and the accumulated firm-specific experience effects, represent human capital that workers have accumulated and can deploy at any employer. While the firm growth components ($\hat{\phi}_{\kappa(j)}$) are acquired through experience at specific employers, once obtained, these skills become part of the worker's portable human capital stock. A worker who spent years at a high-growth firm carries the accumulated benefits of that experience to subsequent employers, making this component fundamentally part of their individual productivity rather than a characteristic of their current employment relationship.

With this dynamic worker effect in hand, we decompose the variance of wages into three time-varying components:

$$\begin{aligned}\text{Var}\left(\ln y_{i,t}\right) = &\underbrace{\text{Var}\left(\hat{\psi}_{J(i,t)}\right)}_{\text{Var(Firm FE)}} + \underbrace{\text{Var}\left(\alpha_{i,t}\right)}_{\text{Var(Dynamic Worker Effect)}} \\ &+ \underbrace{2\,\text{Cov}\left(\alpha_{i,t}, \hat{\psi}_{J(i,t)}\right)}_{\text{2Cov(Dynamic Worker Effect, Firm FE)}} + \text{Residuals}\end{aligned} \tag{15}$$

Figure 11: Dynamic AKM Decomposition



Notes: This figure presents the dynamic AKM decomposition of log wage variance over the life cycle. The decomposition shows: $\text{Var}(\hat{\psi}_j)$ (Firm FE) in red, $\text{Var}(\alpha_{i,t})$ (Dynamic Worker Effect) in blue, $2\,\text{Cov}(\alpha_{i,t}, \hat{\psi}_j)$ (covariance between Dynamic Worker Effect and Firm FE) in orange and residuals in gray. The dynamic worker effect combines the worker fixed effect, worker growth component, and all accumulated firm-specific experience effects, representing the complete portable human capital of each worker at each age.

Figure 11 reveals that the variance of the dynamic worker effect dominates the evolution of wage inequality over the life cycle. This component, which captures all portable aspects of worker productivity, accounts for the overwhelming majority of both the level and growth in wage dispersion. By age 42, the variance of the dynamic worker effect reaches 0.11 log points, representing nearly 70 percent of total wage variance. This dominance underscores a fundamental insight: life-cycle wage inequality in Veneto is primarily about differences between workers rather than differences between firms.

The relatively minor and stable contribution of firm fixed effects to wage variance—hovering around 0.02 log points throughout the life cycle—indicates that firm wage policies, while important for individual workers, play a secondary role in shaping the overall distribution of wages.

Even as workers age and potentially move to different segments of the firm wage distribution, the variance attributable to firm heterogeneity remains remarkably constant. This stability suggests that the firm component of wages operates largely orthogonally to the dramatic changes occurring in the worker component over the life cycle.

Perhaps most intriguingly, Figure 11 documents a substantial increase in the covariance between the dynamic worker effect and firm fixed effects as workers age. This covariance, starting near zero at age 25, grows to contribute 0.02 log points to total variance by age 42 — a magnitude comparable to the entire contribution of firm fixed effects themselves. This pattern reveals that life-cycle sorting between workers and firms is not merely about initial matching but represents a dynamic process that unfolds over entire careers.

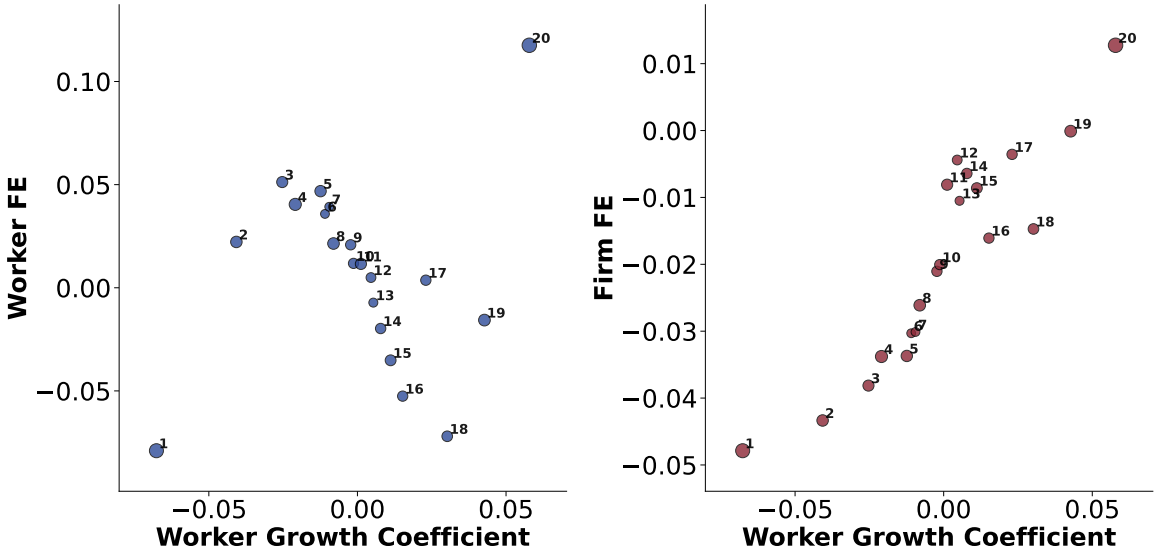
The strengthening correlation between portable worker productivity and firm wage premia reflects the cumulative advantages that accrue to high-ability workers over time. Workers with steep growth trajectories not only see their wages rise through direct experience effects but also are able to access better-paying firms. This growing assortative matching over the life cycle represents a crucial mechanism through which initial differences in growth potential translate into substantial wage inequality by mid-career. The fact that this sorting intensifies with age suggests that the labor market becomes increasingly segmented over the life cycle, with high-growth workers and high-wage firms forming an increasingly exclusive partnership that reinforces and amplifies underlying productivity differences.

7.4 Dynamic Worker-Firm Sorting

We now examine the empirical relationships between worker fixed effects, firm fixed effects, and heterogeneous returns to experience. This analysis reveals how wage levels and growth rates are distributed across workers and firms in our estimated model.

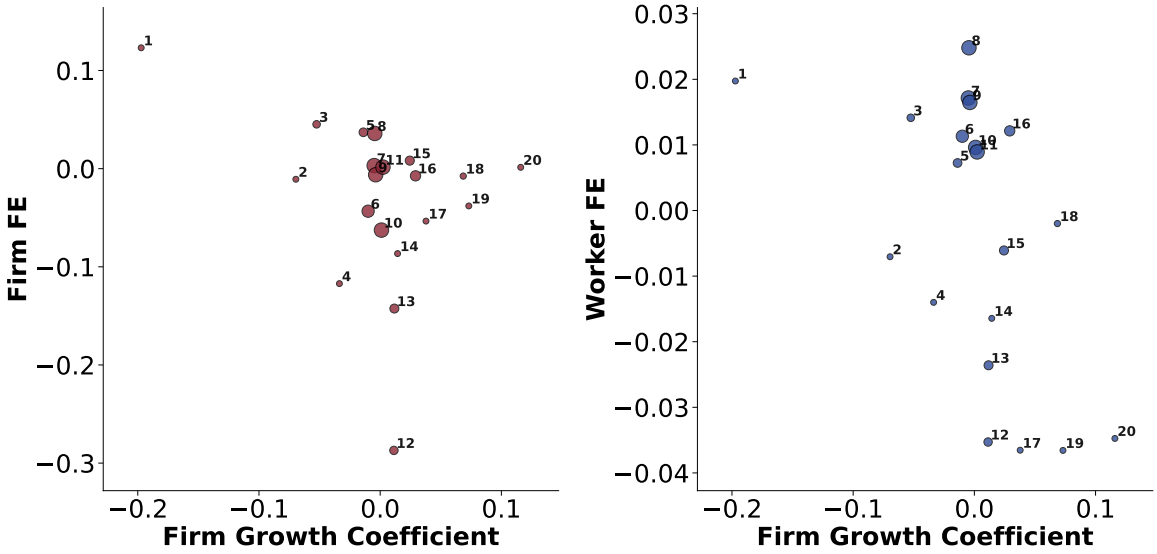
Figure 12 documents two key empirical relationships regarding worker-specific returns to experience. The left panel reveals a negative correlation between worker fixed effects and worker growth coefficients: workers in groups with higher growth coefficients tend to have lower initial worker fixed effects. Specifically, workers in group 1 (lowest growth) have an average worker fixed effect of approximately -0.08, while workers in group 20 (highest growth) have an average worker fixed effect of approximately 0.12. This negative relationship indicates that workers who experience steeper wage growth over their careers tend to start from lower initial wage levels. The right panel 12 shows a contrasting positive relationship between worker growth coefficients and firm fixed effects. Workers with higher growth coefficients are increasingly employed at firms with higher wage premia. Workers in the lowest growth groups (1-5) are employed at firms with average firm fixed effects ranging from -0.05 to -0.03, while workers in the highest

Figure 12: High Wage Workers, High Wage Firms and High Growth Workers



Notes: This figure displays the relationship between worker growth coefficients, worker FE, and Firm FE across the 20 worker groups. The left panel plots the average worker fixed effect $\hat{\alpha}_i$ for each worker growth group against the group's worker growth coefficient $\hat{\beta}_{\rho(i)}$. The right panel plots the average firm fixed effect $\hat{\psi}_j$ for workers (at age 25) in each worker growth group against the group's worker growth coefficient. Groups are numbered 1-20, where group 1 has the lowest worker growth coefficient and group 20 has the highest. Circle size reflects the share of workers in each group.

Figure 13: High Wage Workers, High Wage Firms and High Growth Firms



Notes: This figure displays the relationship between firm growth coefficients, Firm FE, and worker FE across the 20 firm groups. The left panel plots the average firm fixed effect $\hat{\psi}_j$ for each firm growth group against the group's firm growth coefficient $\hat{\phi}_{\kappa(j)}$. The right panel plots the average worker fixed effect $\hat{\alpha}_i$ for workers employed at firms in each firm growth group against the group's firm growth coefficient. Groups are numbered 1-20, where group 1 has the lowest firm growth coefficient, and group 20 has the highest. Circle size reflects the share of firms in each group.

growth groups (16-20) work at firms with average firm fixed effects between 0.01 and 0.015. This positive correlation demonstrates that high-growth workers sort into high-wage firms.

Figure 13 examines how firm-specific returns to experience relate to other wage components. The left panel shows no systematic relationship between firm fixed effects and firm growth coefficients. Firms across the entire spectrum of growth coefficients — from group 1 with a coefficient of approximately -0.15 to group 20 with a coefficient of approximately 0.10 — display firm fixed effects scattered around zero with no clear pattern. The correlation between these two firm-level components is effectively zero. Similarly, the right panel of Figure 13 reveals no clear relationship between worker fixed effects and firm growth coefficients. Workers employed at firms with different growth profiles show no systematic differences in their average worker fixed effects. The worker fixed effects remain scattered between -0.04 and 0.03 across all firm growth groups without any discernible trend.

These empirical patterns reveal important asymmetries in labor market sorting. While workers with steeper experience profiles systematically match with high-wage firms — a relationship that strengthens over the life cycle as documented in our dynamic decomposition — no such sorting occurs along the firm growth dimension. High-growth firms are neither high-wage nor low-wage employers on average, and they employ workers across the full spectrum of the worker fixed effect distribution.

The contrast between the sorting patterns for worker growth effects versus firm growth effects suggests these two sources of wage growth operate through distinct channels in the labor market. Worker-specific returns to experience are integrated into the broader sorting dynamics between workers and firms, with high-growth workers progressively matching with premium-paying employers. Firm experience effects, by contrast, appear orthogonal to other wage components, contributing to wage dispersion without systematically reinforcing or offsetting other sources of wage differences.

8 Conclusion

Why do some workers experience rapid wage growth over their careers while others see their wages stagnate? This fundamental question has motivated decades of research on wage inequality, yet a critical gap remains: we lack a clear understanding of whether diverging wage trajectories primarily reflect differences in how workers accumulate human capital or differences in the types of firms that employ them. This paper provides both a methodological framework and empirical evidence to decompose the sources of life-cycle wage inequality.

Methodologically, we extend the seminal AKM model to incorporate two-sided heterogeneity

in returns to experience across both workers and firms using a grouped fixed-effects approach. Recognizing that estimating individual-level heterogeneity is infeasible with limited observations per worker and firm, we develop a novel two-way clustering algorithm that jointly classifies workers and firms into discrete growth types while estimating group-specific returns to experience. Applying this framework to matched employer-employee data from Veneto, Italy, we document two primary findings. First, worker heterogeneity in returns to experience dominates firm heterogeneity in shaping life-cycle wage inequality. Second, we find that workers with steep experience profiles progressively sort into firms offering higher wage premiums over their careers. This dynamic sorting amplifies wage inequality through both individual growth advantages and differential access to premium employers.

One striking feature of our results is that firms do not function as stepping stones in our context — firm heterogeneity contributes relatively little to rising wage inequality. Two mechanisms explain this pattern. First, while firm growth effects range from -0.2 to 0.1, most workers are concentrated at firms offering near-average returns to experience, limiting the aggregate impact. Second, substantial worker mobility across firm types averages out firm-specific effects over careers, while worker-specific returns remain portable across all jobs.

Our framework opens several promising research directions. One natural application is decomposing the sources of the college wage premium over the life cycle. Does the premium grow because college graduates accumulate human capital faster, work at firms offering steeper returns to experience, or progressively sort into premium employers? Similarly, our approach could illuminate the sources of racial and gender wage gaps over careers, distinguishing between differences in portable human capital accumulation and differential sorting across firm types. Beyond these substantive applications, our two-way clustering methodology extends to any setting requiring estimation of two-sided heterogeneity with high-dimensional datafrom student-teacher matching in education to buyer-seller relationships in industrial organization. Understanding how heterogeneity on both sides of a match shapes outcomes, and whether one side dominates the other, remains crucial across many domains of economic research.

References

Abowd, John M, Francis Kramarz, and David N Margolis, "High wage workers and high wage firms," *Econometrica*, 1999, *67* (2), 251–333.

_ , Robert H Creecy, and Francis Kramarz, "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data," Longitudinal Employer-Household Dynamics Technical Papers 2002-06, Center for Economic Studies, U.S. Census Bureau Mar 2002.

Altonji, Joseph G and Charles R Pierret, "Employer learning and statistical discrimination," *The Quarterly Journal of Economics*, 2001, *116* (1), 313–350.

Alvarez, Jorge, Felipe Benguria, Niklas Engbom, and Christian Moser, "Firms and the decline in earnings inequality in Brazil," *American Economic Journal: Macroeconomics*, 2018, *10* (1), 149–189.

Arellano-Bover, Jaime and Fernando Saltiel, "Differences in On-the-Job Learning across Firms," *Journal of Labor Economics*, 2024, *42* (2), forthcoming.

Bagger, Jesper, François Fontaine, Fabien Postel-Vinay, and Jean-Marc Robin, "Tenure, experience, human capital, and wages: A tractable equilibrium search model of wage dynamics," *American Economic Review*, 2014, *104* (6), 1551–1596.

Bartolucci, Cristian, Francesco Devicienti, and Ignacio Monz'on, "Identifying sorting in practice," *American Economic Journal: Applied Economics*, 2018, *10* (4), 408–438.

Battisti, Michele, "High wage workers and high wage peers," *Labour Economics*, 2017, *46*, 47–63.

Biasi, Barbara and Heather Sarsons, "Information, confidence, and the gender gap in bargaining," *AEA Papers and Proceedings*, 2021, *111*, 174–78.

Bonhomme, Stephane and Angela Denis, "Bias reduction, Groups, Shrinkage, and Factors in Panel Data," Working Paper 2024.

Bonhomme, Stéphane and Elena Manresa, "Grouped Patterns of Heterogeneity in Panel Data," *Econometrica*, 2015, *83* (3), 1147–1184.

_ , Thibaut Lamadon, and Elena Manresa, "A Distributional Framework for Matched Employer Employee Data," *Econometrica*, 2019, *87* (3), 699–739.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa, "Discretizing Unobserved Heterogeneity," *Econometrica*, 2022, *90* (2), 625–643.

Borovičková, Katarína and Claudia Macaluso, "Heterogeneous Job Ladders," *Journal of Monetary Economics*, 2024, p. 103711.

Bradley, James, Bradley Setzler, and Felix Tintelnot, "The Effects of Foreign Multinationals on Workers and Firms in the United States," *The Quarterly Journal of Economics*, 2023, *138* (3), 1499–1542.

Burdett, Kenneth and Dale T. Mortensen, "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, 1998, *39* (2), 257–273.

Caldwell, Sydnee, Ingrid Haegele, and J"org Heining, "Bargaining and inequality in the labor market," 2024. Working Paper.

Card, David, Ana Rute Cardoso, and Patrick Kline, "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women," *The Quarterly Journal of Economics*, 2016, *131* (2), 633–686.

_ , Francesco Devicienti, and Agata Maida, "Rent-sharing, holdup, and wages: Evidence from matched panel data," *Review of Economic Studies*, 2014, *81* (1), 84–111.

_ , Jörg Heining, and Patrick Kline, "Workplace heterogeneity and the rise of West German wage inequality," *The Quarterly Journal of Economics*, 2013, *128* (3), 967–1015.

Dauth, Wolfgang, Sebastian Findeisen, Enrico Moretti, and Jens Suedekum, "Matching in Cities," *Journal of the European Economic Association*, 2022, *20* (4), 1478–1521.

Deming, David J., "Why Do Wages Grow Faster for Educated Workers?," Working Paper 31373, National Bureau of Economic Research 2023.

Gathmann, Christina and Uta Schönberg, "How General is Human Capital? A Task-Based Approach," *Journal of Labor Economics*, 2010, *28* (1), 1–49.

Gregory, Victoria, "Firms as Learning Environments: Implications for Earnings Dynamics and Job Search," Working Paper 2020-036, Federal Reserve Bank of St. Louis 2020.

_ , Guido Menzio, and David Wiczer, "The Alpha Beta Gamma of the Labor Market," *Journal of Monetary Economics*, 2024, p. 103695.

Guvenen, Fatih, "An empirical investigation of labor income processes," *Review of Economic Dynamics*, 2009, *12* (1), 58–79.

_ , Fatih Karahan, Serdar Ozkan, and Jae Song, "What do Data on Millions of US Workers Reveal about Lifecycle Earnings Dynamics?," *Econometrica*, 2021, *89* (5), 2303–2339.

Helm, Ines, Alice Kügler, and Uta Schönberg, "Displacement Effects in Manufacturing and Structural Change," 2023. Working Paper.

Huggett, Mark, Gustavo Ventura, and Amir Yaron, "Sources of Lifetime Inequality," *American Economic Review*, 2011, *101* (7), 2923–2954.

Kline, Patrick, Raffaele Saggio, and Mikkel Sølvsten, "Leave-out estimation of variance components," *Econometrica*, 2020, *88* (5), 1859–1898.

Lentz, Rasmus, Suphanit Piyapromdee, and Jean-Marc Robin, "The Anatomy of SortingEvidence From Danish Data," *Econometrica*, 2023, *91* (6), 2409–2455.

Ma, Xiao, Alejandro Nakab, and Daniela Vidart, "How Do Workers Learn?: Theory and Evidence on the Roots of Lifecycle Human Capital Accumulation," Working Paper, University of Connecticut, Department of Economics 2022.

_ , _ , and _ , "Human capital investment and development: The role of on-the-job training," *Journal of Political Economy Macroeconomics*, 2024, *2* (1), 107–148.

Magnac, Thierry, Nicolas Pistolesi, and Sébastien Roux, "Post-Schooling Human Capital Investments and the Life Cycle of Earnings," *Journal of Political Economy*, 2018, *126* (3), 1219–1249.

Mann, Lukas, "Spatial Sorting and the Rise of Geographic Inequality," February 2024. Job market paper, Department of Economics, Princeton University.

Ozkan, Serdar, Jae Song, and Fatih Karahan, "Anatomy of Lifetime Earnings Inequality: Heterogeneity in Job-Ladder Risk versus Human Capital," *Journal of Political Economy Macroeconomics*, 2023, *1* (3), 506–550.

Postel-Vinay, Fabien and Jean-Marc Robin, "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, 2002, *70* (6), 2295–2350.

Roussille, Nina, "The role of the ask gap in gender pay inequality," *The Quarterly Journal of Economics*, 2024, p. qjae004.

Rubinstein, Yona and Yoram Weiss, "Post Schooling Wage Growth: Investment, Search and Learning," in Eric A. Hanushek and Finis Welch, eds., *Handbook of the Economics of Education*, Vol. 1, Elsevier, 2006, pp. 1–67.

Serafinelli, Michel, ""Good" firms, worker flows, and local productivity," *Journal of Labor Economics*, 2019, *37* (3), 747–792.

Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter, "Firming up inequality," *The Quarterly Journal of Economics*, 2019, *134* (1), 1–50.

Sørensen, Kenneth Lykke and Rune Vejlin, "Worker and Firm Heterogeneity in Wage Growth: An AKM Approach," *Labour*, 2011, *25* (4), 485–507.

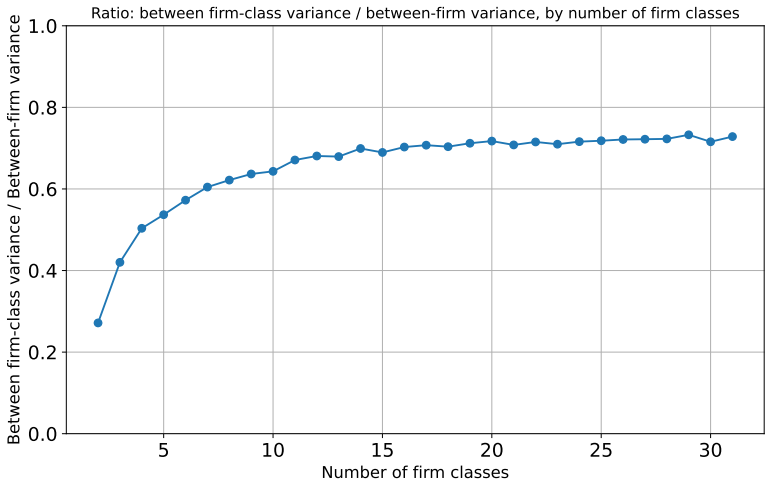
APPENDICES FOR ONLINE PUBLICATION

The Worker and Firm Origins of Life-Cycle Wage Inequality

by Xincheng Qiu, Jesse Wedewer, Runling Wu

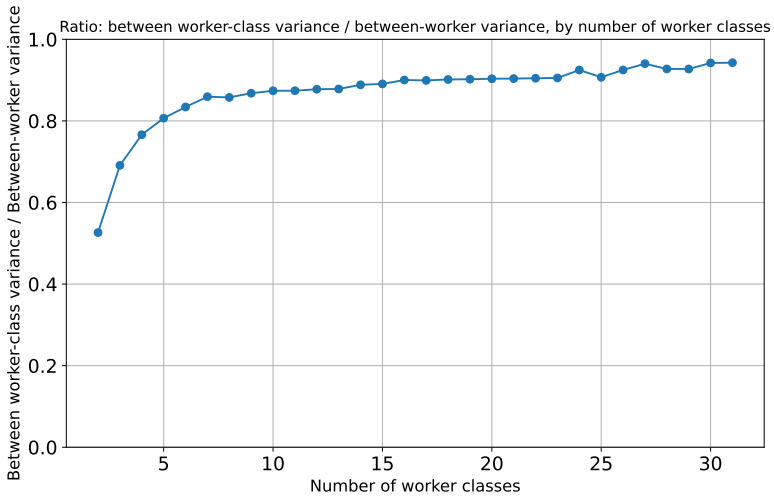
I Additional Figures and Tables

Figure A-1: Number of firm classes - Veneto



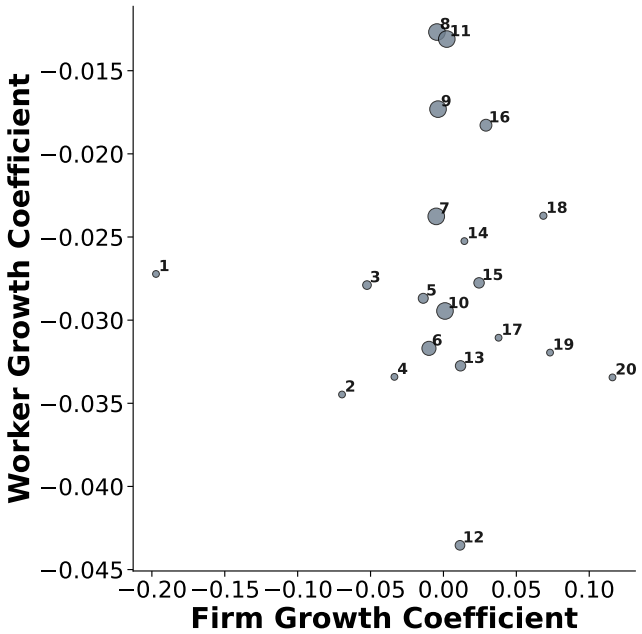
Notes: This figure shows the ratio of between-firm-class variance to total between-firm variance as a function of the number of firm classes.

Figure A-2: Number of worker classes - Veneto



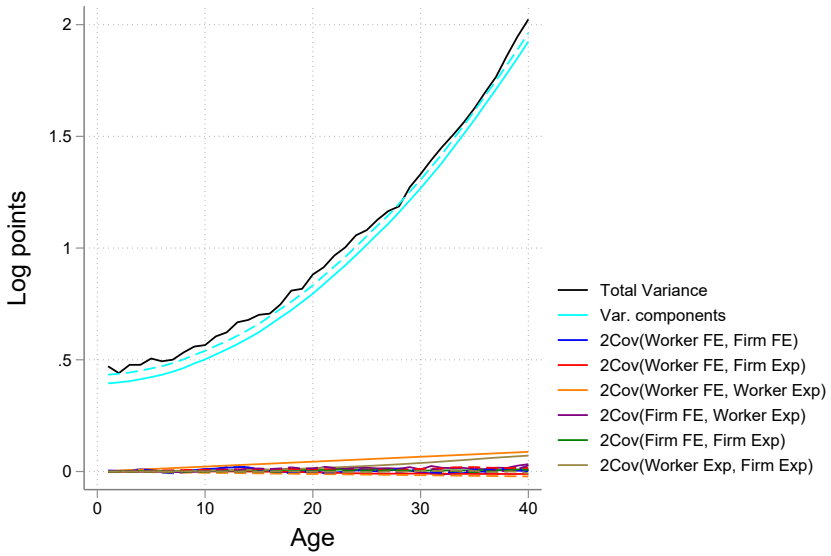
Notes: This figure shows the ratio of between-worker-class variance to total between-worker variance as a function of the number of worker classes.

Figure A-3: High Growth Workers and High Growth Firms



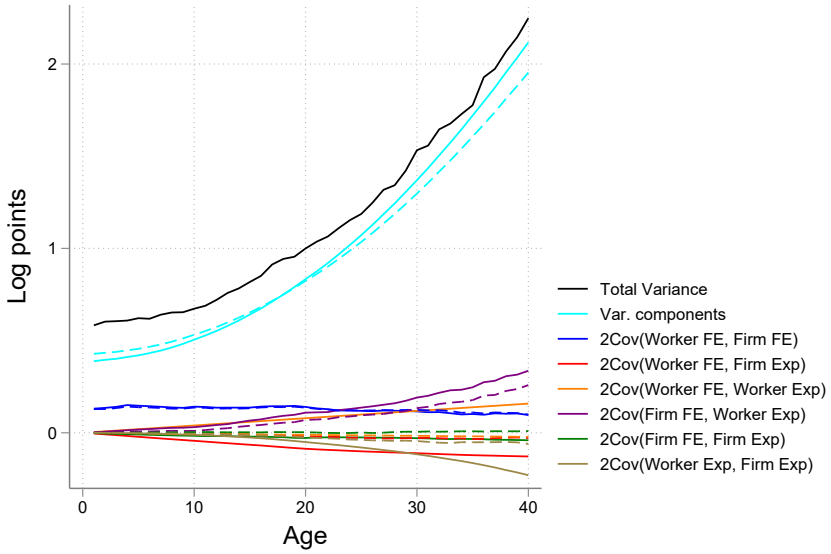
Notes: This figure displays the relationship between worker growth coefficients and firm growth coefficients across the 20 worker groups and 20 firm groups. Groups are numbered 1-20. Circle size reflects the share of workers in each group.

Figure A-4: Variance decomposition over the life-cycle: covariance components



Notes: This figure decomposes the total variance into its covariance components over the life cycle. The black line shows total variance, with the cyan line representing the sum of variance components. Colored lines show twice the covariance between different components: Worker FE and Firm FE (blue), Worker FE and Firm Exp (red), Worker FE and Worker Exp (orange), Firm FE and Worker Exp (purple), Firm FE and Firm Exp (green), and Worker Exp and Firm Exp (brown). Solid lines represent estimates from our algorithm while dashed lines show true values. The minimal covariance between worker and firm components reflects the absence of sorting in this DGP.

Figure A-5: Variance decomposition over the life-cycle: covariance components



Notes: This figure shows covariance components under sorting. The black line represents total variance while colored lines show twice the covariance between components. Notable features include: positive and increasing 2Cov(Worker FE, Firm FE) in blue reflecting assortative matching; growing 2Cov(Firm FE, Worker Exp) in purple as high-growth workers sort to better firms; and negative 2Cov(Worker FE, Firm Exp) in red early in careers that turns positive as sorting strengthens. Solid lines show estimates and dashed lines true values. The algorithm accurately captures these complex sorting dynamics, including the non-monotonic patterns that arise from the interaction of heterogeneity and dynamic matching.

II Assessing Performance of the Algorithm Using Unit-Level Fixed Effects

To assess the validity of our algorithm's performance, we consider an alternative estimation procedure to obtain unit-level fixed effects $\beta_i, \phi_{j(i,t)}$ for a subset of workers and firms. The key idea is that for job stayers, the first difference in log wages may be expressed as an additively separable function of the worker growth coefficient β_i and the firm growth coefficient $\phi_{j(i,t)}$, given that observable controls $X_{i,t}$ are not time-varying. This allows for the application of the framework of Abowd et al. (1999) to assess two-sided unobserved heterogeneity in wage growth to obtain unit-level estimates $\hat{\beta}_i$ and $\hat{\phi}_j$.

Formally, for individuals where $j(i,t) = j(i,t-1)$:

$$\Delta ln(\tilde{y}_{i,t}) = \beta_i + \phi_{j(i,t)} + \varepsilon_{i,t} \tag{A1}$$

where $\Delta ln(\tilde{y}_{i,t}) := ln(\tilde{y}_{i,t}) - ln(\tilde{y}_{i,t-1})$ and $\varepsilon_{i,t} = r_{i,t} - r_{i,t-1}$ [15]. This expression may be estimated for the largest connected set of workers and firms, subject to an arbitrary normalization for one firm $j \in J$ and the identifying assumptions of Abowd et al. (1999).

While the above unit-level two-way fixed effect framework is both natural and intuitive, there several important drawbacks. First, the restriction to observations where workers remained at the same employer for two consecutive years is not innocuous: this not only dramatically reduces the size of the sample but also changes the sample composition by removing short-lived jobs, as shown in descriptive statistics in Table A-1, Appendix II.

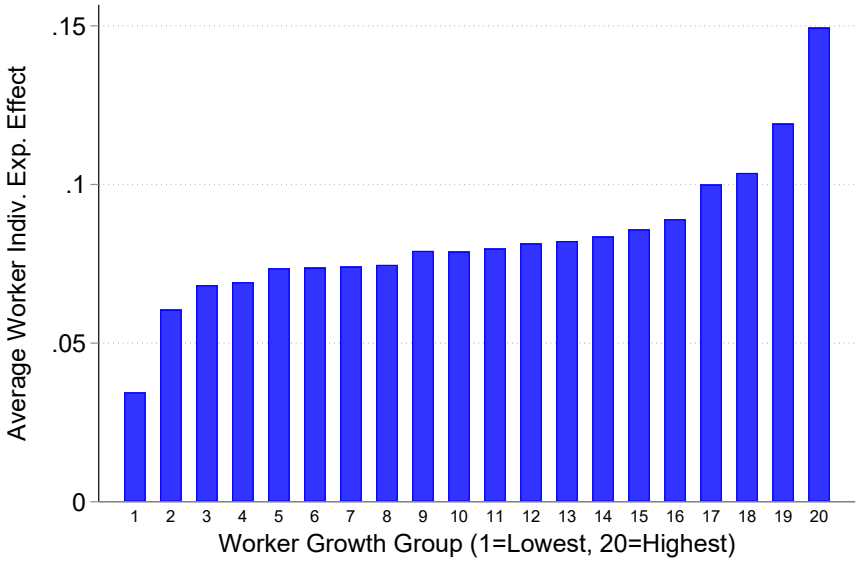
To assess the performance of our two-way clustering algorithm, we compare our group-level estimates with individual-level fixed effects estimates obtained from the alternative estimation procedure described above. Figures A-6 and A-7 present the average individual-level growth effects within each worker and firm growth group, respectively. For worker groups, we observe a nearly monotonic relationship between group assignments and average individual effects. Workers assigned to higher growth groups exhibit substantially higher average individual-level returns to experience, with the relationship particularly steep for the highest growth groups. The pattern for firm groups is also generally increasing, though with somewhat greater variation, partially reflecting the additional computational challenges in firm assignment discussed in Section 3.

Figures A-8 and A-9 in Appendix II plot the group-level estimates against the average individual effects within each group, with both normalized to set the minimum value to zero to

[15] Given that in our context, $X_{i,t}$ are gender-specific age effects (quadratic), we work with residualized log daily wages where $\tilde{y}_{i,t}$ is derived from a regression $ln(y_{i,t}) = \alpha_0 + \xi X_{i,t} + ln(\tilde{y}_{i,t})$.

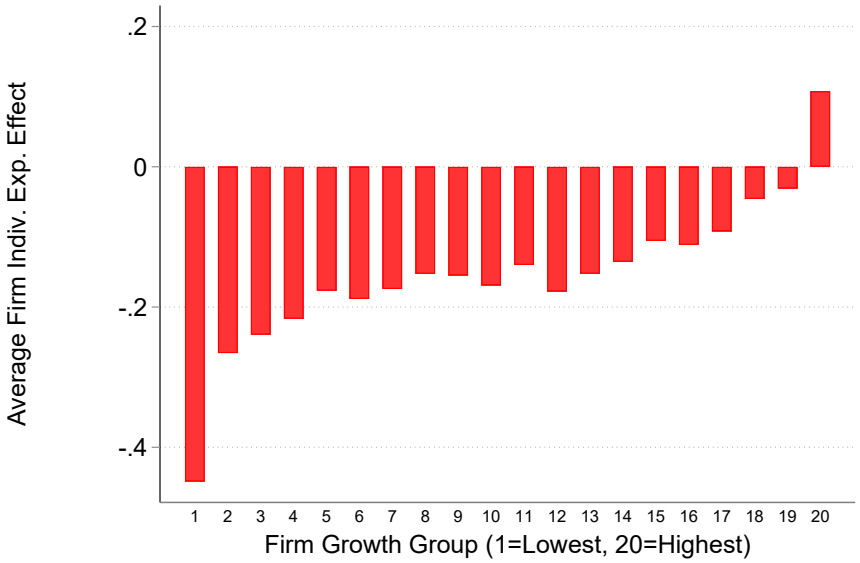
ensure comparability given the sensitivity to firm normalization. The scatter plots reveal a strong positive relationship for both workers and firms, with points lying close to the 45-degree line. This alignment indicates that our clustering algorithm successfully groups workers and firms with similar growth characteristics.

Figure A-6: Average Unit-Level Worker Growth Effect by Worker Group



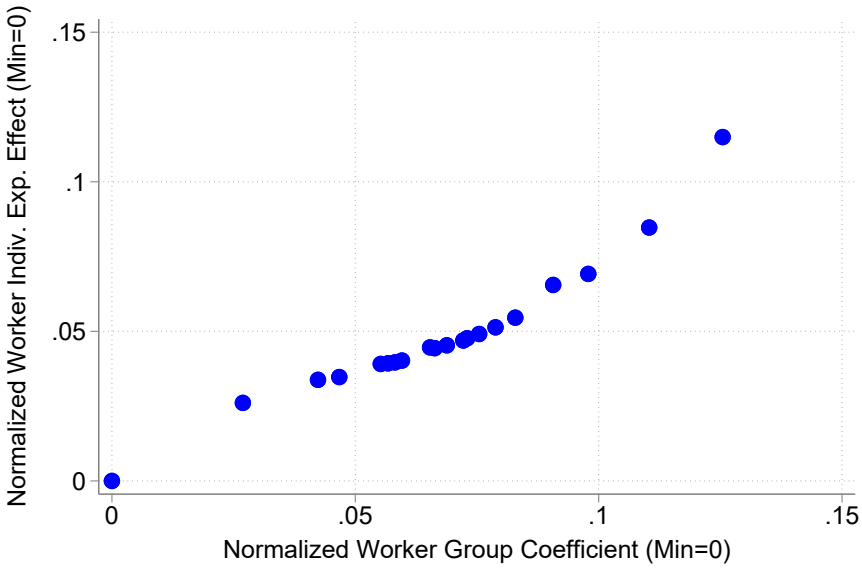
Notes: This figure plots the estimated conditional average fixed effect for each worker growth group, $\bar{\beta}_i|l = \sum_i 1[\rho(i) = l]\bar{\beta}_i$, weighted at the person-year level. Worker growth groups are labeled according to the magnitude of $\hat{\beta}_{\rho(i)}$, such that the worker growth group 1 has the lowest value and the worker growth group 20 has the highest value.

Figure A-7: Average Unit-Level Firm Growth Effect by Firm Group



Notes: This figure plots the estimated conditional average fixed effect for each firm growth group, $\bar{\phi}_j|k = \sum_j 1[\kappa(j) = k]\bar{\phi}_j$, weighted at the person-year level. Firm growth groups are labeled according to the magnitude of $\hat{\phi}_{\kappa(j)}$, such that firm growth group 1 has the lowest value and firm growth group 20 has the highest value.

Figure A-8: Average Unit-Level Worker Growth Effect by Worker Growth Group



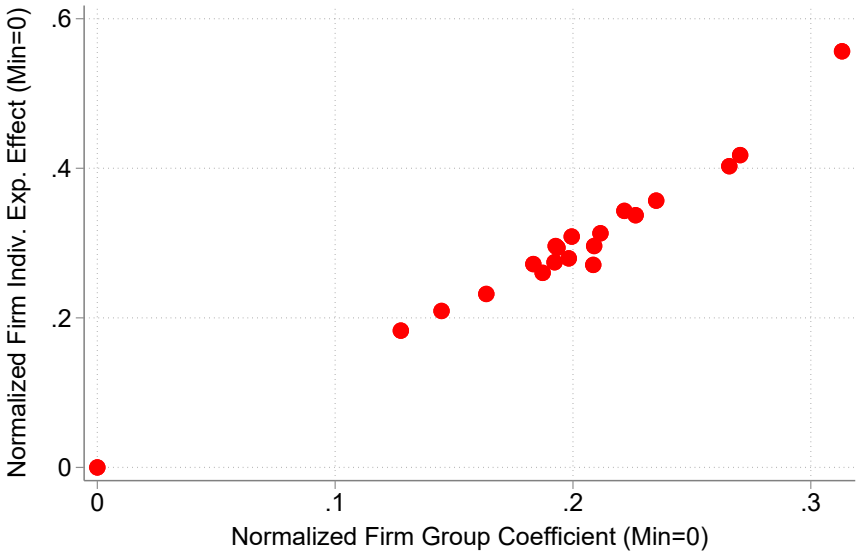
Notes: This figure plots the conditional average fixed effect for each worker growth group, $\bar{\beta}_i|l = \sum_i 1[\rho(i) = l]\bar{\beta}_i$ against the group-level estimate, $\hat{\beta}_{\rho(i)}$, weighted at the person-year level. Importantly, the conditional average fixed effect is normalized such that the lowest group-level average takes on a value of 0. As estimates are sensitive to firm normalization, such a normalization is important to ensure comparability of estimates.

Table A-1: Summary Statistics of Job Stayers

	(1) Mean	(2) SD
Worker Characteristics		
Fraction of male	0.66	0.47
Worker's age	31.1	3.97
Tenure since age 25	4.12	3.02
Experience since age 25	5.30	3.52
Job, Wage Characteristics		
Primary job daily wages	134.36	56.58
Job duration (in days)	237.26	104.19
Annual wage growth	0.032	0.158
Firm Characteristics		
Firm size	9.03	69.86
Number of person-year observations	4,588,997	
Number of person-year observations in largest connected set	3,226,863	
Number of unique workers	879,386	
Number of unique firms	263,104	
Number of unique firms in largest connected set	60,202	
Number of 'moves' in largest connected set	211,732	

Notes: The table reports means and standard deviations of each variable based on the Veneto Worker History Panel for observations of job stayers where daily wages are observed in two consecutive periods.

Figure A-9: Average Unit-Level Firm Growth Effect by Firm Growth Group



Notes: This figure plots the conditional average fixed effect for each firm growth group, $\bar{\phi}_j|k = \sum_j 1[\kappa(j) = k]\bar{\phi}_j$ against the group-level estimate, $\hat{\phi}_{\kappa(j)}$, , weighted at the person-year level. Importantly, the conditional average fixed effect is normalized such that the lowest group-level average takes on a value of 0. As estimates are sensitive to firm normalization, such a normalization is important to ensure comparability of estimates.