

Heterogeneous Skill Portability across Firms

Job Market Paper^{*}

Eugenio Renedo[†]

CEMFI

This version: October 24, 2026

Updated regularly; please [click here](#) for the latest version.

Skills learned on the job vary in their portability across employers. Some are portable and retain their value when workers change jobs, while others are firm-specific and are only productive at the employer where they were acquired. This paper studies how differences across firms in the portability of skills learned on the job shape wage inequality within occupations and wage losses following displacement. I develop an equilibrium search model in which firms differ in the extent to which they equip workers with portable versus firm-specific human capital. I estimate the model using French administrative data that link workers' careers to firm balance sheets and closures. The estimated share of on-the-job learning that is firm-specific averages 55%, with substantial heterogeneity across firms. Counterfactual simulations show that equalizing this share across firms at its baseline mean reduces within-occupation variance in log wages by 18%. This decline is concentrated among workers who enter the labor market with fewer skills and whose wages depend more on human capital accumulated on the job. Eliminating cross-firm differences in the portability of learning also reduces average wage losses two to four years after displacement by 23%.

Keywords: human capital, portability, search frictions, wage inequality, job displacement.

JEL codes: E24, J24, J31, J62, J63.

^{*}I am indebted to Manuel Arellano, Nezih Guner and Kurt Mitman for invaluable guidance and support. I also thank Jaime Arellano-Bover, Anmol Bhandari, Alexander Bick, Stéphane Bonhomme, Christine Braun, Tomás Budí-Ors, Sebastián Fanelli, Victoria Gregory, Philipp Grübener, Loukas Karabarbounis, Federico Kochen, Julian Kozłowski, Joseph Mullins, Serdar Ozkan, Josep Pijoan-Mas, B. Ravikumar, and Guillaume Vandenbroucke, as well as participants at the St. Louis Fed PhD Dissertation Fellows Workshop 2026, the XXIX Workshop on Dynamic Macroeconomics 2026, and the CEMFI Macroeconomics Workshop for valuable comments and discussions. I gratefully acknowledge financial support from grants PREP2022-000905 (funded by MICIU/AEI/10.13039/501100011033 and ESF+), MDM-2016-0684, and CEX2020-001104-M (both funded by MICIU/AEI/10.13039/501100011033). All errors are my own.

[†]CEMFI, Casado del Alisal 5, 28014 Madrid, Spain. Email: eugenio.renedo@cemfi.edu.es. Website: erenedo.github.io

1. Introduction

Firms are learning environments where workers build skills on the job. However, not everything workers learn travels with them when they change employers. Some skills are *portable* and remain valuable across firms, while others are *firm-specific* and lose their value when the employment relationship ends. Importantly, workers in the same occupation may acquire different amounts of portable and firm-specific skills depending on the firm they work for. If so, these differences may have major implications for workers' careers, shaping both wage growth and the wage losses associated with job displacement. This paper asks two questions. Within occupations, do firms differ in the portability of skills they build? And, do these differences contribute to within-occupation wage inequality and wage losses following job displacement?

To answer these questions I develop an equilibrium search model in which workers accumulate portable and firm-specific human capital, and firms differ in how intensively they use and build each type. I estimate the model on French administrative data and use it to measure how the mix of portable and firm-specific learning varies across jobs within the same occupation, something that cannot be directly observed in the data. I find that, on average, 55% of on-the-job skill learning is firm-specific, with this share varying across firms from 23% at the tenth percentile to 86% at the ninetieth percentile. Finally, I study a counterfactual economy in which all firms use and build the two types of human capital in the same proportion. Comparing this economy with the baseline reveals that, absent these differences in the share of firm-specific learning across firms, the variance of within-occupation log wages would be 18% lower. The reduction is concentrated among workers who enter the labor market with fewer skills and whose wages depend more on human capital accumulated on the job. Likewise, removing differences across firms in the portability of skills learned on the job would reduce the average wage loss from displacement by 23%, with the reduction being particularly pronounced for long-tenured workers who tend to have accumulated more firm-specific human capital.

Measuring differences in the portability of skills across firms poses a fundamental empirical challenge: neither the human capital workers accumulate on the job nor its portability across employers is directly observable. Both must be inferred from their effects on wages and output, making it essential to distinguish their empirical implications from those of other factors that shape these outcomes.

The model is designed to address this challenge. Workers accumulate both portable and firm-specific human capital while employed, and firms' technologies differ in the intensity with which they use each type of human capital in production. Because workers learn what they use at work, technological differences across firms shape the portability of the skills acquired on the job. These human capital gains raise firm output, and workers and firms bargain over the resulting additional surplus. To identify the effects of cross-firm differences in skill portability, the model also incorporates other mechanisms that shape wages and output in the data. Workers receive outside offers that allow them to move up the job ladder, affecting both wage growth and learning opportunities. Firms, meanwhile, experience

productivity shocks that can lead to match dissolution. When a match ends, workers fall off the job ladder, making job loss costly.

I estimate the model using French administrative data from 2008 to 2016. The main dataset follows a large, representative sample of private-sector workers throughout their employment and unemployment histories. I link these records to firms' balance sheets and information on firm closures. Combining worker and firm data is essential for identifying the model's different components. Wages upon entering new jobs provide information about portable human capital. Less experienced workers may accept lower wages when hired into firms that build mainly portable skills that are valuable across firms. Conversely, more experienced workers command higher wages when entering these jobs because they already possess portable skills and have less to gain from further learning. The relationship between wages upon entering a new job and prior labor market experience identifies the intensity with which firms use portable human capital. Firm-specific human capital, by contrast, is valued only by the employer that built it and is lost when the match ends. Workers with longer tenure tend to have accumulated more firm-specific human capital, so the wage losses they suffer when displaced by a firm closure are informative of the intensity with which firms use firm-specific human capital. Finally, the evolution of firms' value added identifies productivity dynamics independently of wages, and job flows pin down the labor market environment and the job ladder.

I validate the model against two facts in the data that were not targeted in the estimation. The facts are closely linked to the portability of skills learned on the job and the workers' labor-market outcomes I want to study. I show that reproducing these moments requires explicitly modeling firm-specific human capital and cross-firm differences in skill portability. To establish this, I compare the model's predictions with those of a restricted version that abstracts from firm-specific human capital but is otherwise identical.

The first fact compares incumbent workers, who were already employed at the firm in the previous year, with newcomers, who joined during the year. Incumbents contribute 5.6 log points more to firm output than newcomers, but earn only 4.3 log points more in wages. Thus, 1.3 log points of incumbents' output advantage are not reflected in their wages. The estimated model reproduces this gap, with a predicted wedge of 1.6 log points, close to the empirical estimate, whereas the restricted model predicts a wedge of only 0.4 log points. This pattern is consistent with the model's mechanism: workers accumulate firm-specific skills on the job that are not valued by other employers, limiting their ability to capture the full value of their contribution to output. Because the restricted model does not include firm-specific human capital, it cannot adequately account for the gap between incumbents' contributions to output and their wages.

The second untargeted fact examines the relationship between the returns to tenure and experience across jobs within the same occupation. To estimate these returns, I use a grouped fixed-effects estimator in the spirit of Bonhomme and Manresa (2015), which classifies workers and jobs into latent

classes. The correlation between the two returns across job classes is 0.25 in the data and 0.30 in the estimated model, compared with 0.80 in the restricted model. The model explains this pattern through differences in the portability of human capital. In the restricted model, all human capital is portable, so jobs differ only in how intensively they use portable human capital. Jobs that use it more intensively offer higher returns to workers' accumulated experience. They also generate faster wage growth with tenure, because the skills workers acquire on the job are valued by all employers. Consequently, the returns to experience and tenure move together, generating a strong positive correlation across jobs. In the estimated model, by contrast, jobs can differ in their use of firm-specific human capital. Jobs that rely more heavily on this type of capital offer higher returns to tenure without necessarily offering higher returns to experience, because the skills workers accumulate are valuable only to their current employer. This weakens the relationship between the two returns, bringing their correlation closer to the empirical estimate.

The estimated model reveals substantial heterogeneity across firms in the portability of learning. On average, 55% of on-the-job learning is firm-specific, but this share ranges from 23% at the tenth percentile to 86% at the ninetieth percentile of firms. Thus, workers in the same occupation but employed at different firms may accumulate very different amounts of portable and firm-specific human capital. To quantify the implications of these differences for within-occupation wage inequality and wage losses following displacement, I construct a counterfactual economy in which all firms use and build portable and firm-specific human capital in the same proportion, fixing the firm-specific share at the baseline mean.

The quantitative analysis shows that differences in the portability of learning across firms are an important source of wage inequality, particularly at the bottom of the wage distribution. Equalizing the mix of portable and firm-specific learning across firms reduces the variance of log wages by 18%. The bottom 5% of the wage distribution accounts for 46% of this reduction, while the variance of log wages among workers with low pre-market skills falls by 41%. These workers rely heavily on human capital accumulated on the job, making their wages particularly sensitive to cross-firm differences in the portability of acquired skills. By contrast, the wages of workers with high pre-market skills depend primarily on the skills they bring to the labor market and are therefore less sensitive to differences in the portability of learning across firms.

Similarly, eliminating differences in learning portability across firms reduces average wage losses following displacement by 23%, even though the economy-wide average mix of portable and firm-specific learning remains unchanged. Firms that rely more heavily on firm-specific human capital also induce workers to accumulate more of it. Workers at these firms therefore face a double loss when displaced: they lose a larger stock of firm-specific skills, and those skills were especially valuable in the jobs they leave behind. This reinforcing relationship magnifies the contribution of firm-specific human capital to output, and hence the wage losses when those skills are destroyed. Equalizing the learning mix across firms breaks this relationship, reducing the value of human capital lost upon

displacement. The effects are particularly pronounced for long-tenured workers: wage losses fall by 20% for those with one to three years of tenure and by 26% for those with ten or more years, flattening the rise of displacement losses with tenure by about 36%.

Related literature. My paper is most strongly connected to the structural literature on human capital accumulation in frictional labor markets. Existing models differ primarily along two dimensions: whether human capital is partly firm-specific and whether employers differ in the learning opportunities they provide. Bagger et al. (2014) combine on-the-job search and sequential-auction wage setting with a single, fully portable stock of human capital that accumulates independently of the employer (see also Yamaguchi 2010). Similarly, Jung and Kuhn (2019) and Burdett, Carrillo-Tudela, and Coles (2020) allow general human capital to accumulate on the job and depreciate during unemployment (see also Karahan, Ozkan, and Song 2019; Engbom 2022). Lise and Postel-Vinay (2020) and Adda and Dustmann (2023) enrich the skill space by tying accumulation to occupations or task sectors, but learning remains employer-independent and fully portable. Gregory (2021) and Borovičková and Macaluso (2025) allow employers to differ in how much human capital they generate (see also Balke and Lamadon 2022), but all accumulated human capital remains portable. Jarosch, Oberfield, and Rossi-Hansberg (2021); Herkenhoff et al. (2024); Adenbaum, Babalievsky, and Jungerman (2024) locate this employer heterogeneity in coworker composition rather than a firm-specific learning technology. By contrast, Audoly, De Pace, and Fella (2022) model both general and firm-specific human capital, but all employers accumulate each type at the same rate and job offers are drawn from an exogenous distribution, precluding equilibrium counterfactuals. My paper is the first to combine employer heterogeneity in both the amount and the portability of learning within an equilibrium framework, allowing counterfactual analysis of how heterogeneity in the mix of portable and firm-specific learning affects workers' labor market outcomes.

This paper also contributes to the empirical literature on the wage returns to tenure and experience (Altonji and Shakotko 1987; Topel 1991; Margolis 1996; Altonji and Williams 2005; Dustmann and Meghir 2005; Beffy et al. 2006; Bowlus and Liu 2013; Rubinstein and Weiss 2006; Abowd, Kramarz, and Roux 2006; Buchinsky et al. 2010; Pires, Szydlowski, and Zhao 2024; Arellano-Bover and Saltiel 2026). Theodoropoulos and Singleton (2026) survey forty-seven studies across thirteen countries and conclude that “imposing common returns to tenure across firms and workers risks masking heterogeneity and generating misleading average effects”. A small set of papers measures that heterogeneity directly. Abowd, Kramarz, and Margolis (1999) already observed that seniority effects are “quite heterogeneous across firms”; Margolis (1996) shows that allowing firm-specific compensation policies drives the mean return to zero while leaving a wide cross-firm dispersion; Abowd, Kramarz, and Roux (2006) find a fifth of French firms with significantly positive returns to seniority and a fifth with significantly negative ones; and Pires, Szydlowski, and Zhao (2024) estimate firm-specific returns to both margins. Closest to my paper is Arellano-Bover and Saltiel (2026), who sort firms into learning classes by the wage growth of their stayers and find wide heterogeneity in returns to experience across firms. My contribution is to

measure the *joint* distribution of the returns to tenure and experience across jobs within occupations, at the level of the firm–occupation cell, using a grouped fixed-effects estimator with latent classes of both workers and jobs in the spirit of Bonhomme and Manresa (2015).

Lastly, this paper speaks to the literature on the cost of job loss, which since Jacobson, LaLonde, and Sullivan (1993) has sought the sources of the displacement scar: the lost worker–employer component (Lachowska, Mas, and Woodbury 2020), the loss of secure jobs (Krolikowski 2017; Jarosch 2023), reemployment at lower-skill occupations (Huckfeldt 2022), the specialization of the worker’s knowledge (Hernández-Martínez, Holter, and Pinheiro 2025), and the length of the intervening jobless spell (Fallick et al. 2025). How much the employer margin accounts for is unsettled: estimates of its share run from under a fifth (Burdett, Carrillo-Tudela, and Coles 2020) to about half (Audoly, De Pace, and Fella 2022) to most of it (Jung and Kuhn 2019; Schmieder, von Wachter, and Heining 2023). In all of these accounts the specificity resides in the worker’s occupation, in the specialization of her skills, or in the quality of the job she lost. My contribution is to locate specificity in the technology of the employer the worker loses, and to ask how much of the wage cost of job loss comes from differences across jobs in the mix of portable and firm-specific learning.

Outline. The remainder of the paper is organized as follows. Section 2 describes the three administrative sources combined in the analysis. Section 3 introduces the structural model. Section 4 describes the estimation and identification, presents the estimates, and recovers the learning mix across jobs. Section 5 confronts the estimated model and a restricted model without firm-specific human capital with two untargeted facts on output and wage growth. Section 6 uses a counterfactual economy to quantify the contribution of differences in the learning mix across jobs to wage dispersion and to the wage cost of job loss. Section 7 concludes.

2. Data: Careers, Balance Sheets, and Closures

This section introduces the three French administrative data sources used throughout the paper. First, a matched employer–employee panel provides employment and unemployment histories for a large sample of workers. Second, firm balance sheets provide firm-level measures of value added, employment, and tangible fixed assets. Third, a near-universe register of employment relationships covers each firm’s complete workforce. The register cannot replace the panel: its worker identifiers cannot be linked from one annual file to the next and it does not record unemployment, so careers cannot be followed in it. Instead, because the entire workforce of each firm is observed, it allows me to identify firm closures and who leaves or joins a firm in a year.

2.1. The worker panel

The core of the data is the panel constructed from the *Déclaration Annuelle des Données Sociales* (DADS), the annual declaration that every French employer is legally required to file for each employee. The panel, constructed by INSEE, France’s national statistical institute, follows a representative sample of workers and records their complete employment histories.¹ For each employment spell I observe gross earnings, hours paid, start and end dates, tenure with the employer, labor-market experience, occupation, and the identifiers of the employing firm and establishment, together with the firm’s industry and size and the worker’s age, gender, and region. From 2008 the panel also records compensated unemployment spells so that transitions between employment and unemployment can be followed.

The unit of observation in the panel is the *spell-year*. A spell is a continuous period during which the worker’s situation does not change: either an uninterrupted employment relationship with one employer, or an uninterrupted period on unemployment benefits. Because the data are reported year by year, each spell is split into annual pieces: one observation for each calendar year in which the spell is active. Two examples fix ideas. An employment relationship running from March 2010 to June 2012 contributes three observations, one each for 2010, 2011, and 2012. A worker who loses her job in April, draws unemployment benefits until August, and starts a new job in September contributes three observations within that single year: an employment spell-year, an unemployment spell-year, and a second employment spell-year.

The sample consists of spells of private-sector workers aged 16 to 65 in mainland France between 2008 and 2016.² The wage is defined as hourly gross pay, calculated by dividing gross pay by the hours paid within the spell. These are the hours for which the employer reports paying the worker, including paid overtime, rather than the contractual hours. The sample includes both open-ended and fixed-term contracts. Fixed-term contracts account for a large share of hires in France, and many become open-ended at the same employer; keeping only open-ended contracts would describe a selected minority of entries and keep only the matches that survived their first contract. Wages are deflated using the consumer price index; all monetary quantities in the paper are expressed in 2015 euros. Panel A of Table 1 describes the resulting sample. Appendix A details the remaining sample-selection steps.

2.2. Jobs, firms, and occupations

I define a *job* as the combination of a firm and an occupation. Three implications of this definition are worth spelling out. First, a job is not a single position: several workers can hold the same job at

¹The panel follows workers born in October of even-numbered years, about 1/25 of the salaried population.

²From 2017, employment records are drawn from the monthly DSN (*Déclaration Sociale Nominative*) rather than the annual DADS. Because the hours reported in the DSN are not measured on the same basis as those in the DADS, hourly wages are not comparable across this break in the data. Since compensated unemployment spells are recorded only from 2008, I restrict the sample to 2008–2016.

the same time. If a firm employs twelve machinists, those twelve workers hold one and the same job. Second, the same occupation at two different firms constitutes two different jobs: a machinist at firm A and a machinist at firm B do the same kind of work, but they hold different jobs. Third, two occupations inside the same firm constitute two different jobs: a firm employing machinists and clerks operates two jobs; a firm employing only machinists operates one. A *firm*, throughout, is the legal entity that owns a balance sheet and can die; a firm operates as many jobs as it employs occupations.

Occupations are defined at the three-digit level of the French *Professions et Catégories Socioprofessionnelles* (PCS) classification. I retain the 108 occupations with at least 10,000 employed spell-year observations over the sample period; the 129 excluded occupations together account for 1.1% of employed spell-years. Placing an employment relationship in a job requires giving it an occupation. In the raw data, each spell-year of a relationship carries its own occupation code, and these codes are not always identical across the years of the same relationship. I therefore assign each relationship a single occupation for its entire duration: the code it reports most often across its spell-years. A code change within a relationship, whether a recoding or a genuine move to another position inside the firm, therefore neither splits the relationship into two jobs nor restarts tenure: tenure throughout is the number of years a worker has spent with her employer, not in her current position. Appendix A reports how often codes change within a relationship and why, and explains why employer tenure, rather than position tenure, is the appropriate measure.

2.3. Firm balance sheets

Firm balance-sheet information comes from the FARE files (*Fichier Approché des Résultats d'Esane*), the annual structural business statistics compiled by INSEE from firms' tax filings. The data cover the universe of firms filing in France, with one observation per firm-year. The DADS and FARE panels are linked by firm identifier and calendar year: each employment spell-year is matched to its employer's balance sheet for that year.

From the balance sheet I take the firm's value added, its tangible fixed assets, and its annual-mean employment in full-time equivalents. I define value added per worker as the ratio of value added to employment, deflated by the same price index applied to wages, and the labor share as the firm's wage bill over its value added. Panel B of Table 1 reports value added per worker, the labor share, and firm size, weighted by employment spell-years.

2.4. Firm closures

A central question of the paper is how much workers lose when they are *exogenously* separated from their jobs. Measuring this cost requires separations that are unrelated to the workers themselves. Workers who quit typically do so because they have found a better opportunity, while workers dismissed for

TABLE 1. Sample description.

	Mean	Median	S.D.
<i>Panel A: Worker panel – DADS</i>			
Hourly wage (euros)	16.65	14.15	7.89
Tenure (years)	5.96	3.29	7.07
Experience (years)	16.29	14.86	10.58
Spell-year observations		9,332,641	
employment		8,148,939	
unemployment		1,183,702	
Workers		1,297,380	
Jobs (firm $\times$ occupation)		1,242,080	
Firms		730,179	
Occupations (3-digit PCS)		108	
<i>Panel B: Firm accounts – FARE</i>			
Value added per worker (thousand euros)	104.4	57.9	172.3
Labor share (wage bill over value added)	0.61	0.59	0.26
Firm size (full-time equivalent employees)	6,837	90	28,890
Firms linked to balance sheets		592,600	
Share of employment spell-years linked		0.80	

NOTES: Private-sector workers aged 16–65 in mainland France, 2008–2016. Panel A: an observation is a worker’s employment or unemployment spell within a calendar year; the hourly wage is gross pay divided by hours paid, over full-time employment spells on standard contracts. Panel B: firm variables are observed at the firm–year level and weighted by employment spell-years. The labor share is the firm’s wage bill over its value added and excludes employer social contributions; value added per worker and the labor share are winsorized at the first and ninety-ninth percentiles. All monetary quantities are in 2015 euros.

cause are selected based on their individual performance. Wage changes following these separations therefore combine the effect of job loss with the reasons the separation occurred. Firm closures, by contrast, displace the entire workforce simultaneously, regardless of any individual worker’s characteristics or performance. They therefore provide plausibly exogenous separations, making them well suited for identifying the causal effect of job loss on wages. Throughout, I measure the cost of job loss by the loss in hourly wages once displaced workers are re-employed, and refer to it as the wage cost of job loss.

Identifying a closure requires seeing a firm’s whole workforce, so I use the *Base Tous Salariés* (BTS), a near-universe register of every employment relationship at every employer in France, about sixty million per year. Its unit of observation is the *poste*: one worker at one establishment in one year. Each annual file covers the current and the previous year, so a firm’s disappearance, and where its workers went, can be read within one file.

A firm is a candidate for closure if it employs workers in one year and none in the next. Many such disappearances are not closures: the firm was taken over, merged, or re-registered under a new identifier, and its workforce keeps working together under a new name. Following Hethey-Maier and Schmieder (2013), I trace each candidate's workers to their employers in the following year. If no single successor employs more than a small share of them, the firm has closed. If one successor employs most of them, the firm changed hands or changed identifier. The cases in between are set aside. Appendix A gives the thresholds, the size floor, and the treatment of workers not re-employed. The procedure identifies 204,080 closures with last active years from 2008 to 2015. A closing firm sheds workers before its last recorded year, and those workers are displaced by the same closure; counting only the workers present at the very end would keep a selected group, the last to leave. Each closure therefore contributes two event years, its last active year and the year before it, which I call the run-up year, and I refer to a firm in either of these years as a closing firm.

2.5. Residualization

Wages and output vary in the data for reasons that are outside the scope of this paper. Men and women are paid differently; pay and productivity differ across regions and industries; the whole economy moves with the business cycle; and some occupations pay more than others. Apart from the descriptive statistics of Table 1, statistics computed from the worker panel are therefore based on variables from which this variation has been removed. Two variables are adjusted, each on characteristics observed at its own level, and the residuals are re-centered at their sample means. The log hourly wage, a worker-level variable, is projected on the worker's gender, the region of her establishment, the two-digit industry of her firm, the year, and her three-digit occupation. Log value added per worker, which comes from the firm's balance sheet, is projected on the firm's two-digit industry and the year.

The wage is projected on occupation because the object of interest is how jobs differ within an occupation: whether two employers in the same line of work build different skills in their workers, and pay differently for them. That surgeons earn more than cashiers is a real difference in pay, but it is between-occupation variation, and no part of the analysis is meant to explain it. Removing occupation-level pay differences from the wage keeps exactly the variation the question is about: how an employer pays relative to other employers in the same line of work. Value added per worker is projected only on firm-level characteristics: a balance sheet is reported for the firm as a whole and belongs to none of its workers or occupations, so projecting it on a characteristic of the individual worker in the row would manufacture differences within the firm that the source data do not contain.

3. The Model

This section presents a continuous-time equilibrium search model of the labor market with heterogeneous workers and firms, on-the-job search, endogenous wage determination, and human capital accumulation. Workers differ ex ante in pre-market skill and ex post through the accumulation of portable and firm-specific human capital. Portable human capital is transferable across employers and remains productive at any firm, whereas firm-specific human capital is lost when a match ends. Each firm employs a single worker; hence, throughout the description of the model, the terms *firm*, *job*, and *employer* are used interchangeably. Firms differ in productivity and in the intensity with which they use each type of human capital in production. Productivity changes stochastically, while the intensities with which a firm uses portable and firm-specific human capital remain fixed.

Workers accumulate human capital on the job through stochastic learning opportunities. Both the size of each learning increment and its allocation between portable and firm-specific human capital depend on the employing firm's usage intensities. Consequently, jobs differ not only in the amount of human capital they generate but also in the type of human capital they produce.

Labor market frictions generate unemployment and vacancies, and the matching technology allows both unemployed and employed workers to search for jobs. Portable human capital is not always fully transferable across matches. With a fixed probability, a meeting is classified as *cross-occupational*, in which case the worker's portable human capital is scaled down by a constant factor if the match forms.³ When workers meet potential employers, wages are determined by bargaining over the match surplus. A match with an unemployed worker forms if its surplus is positive. Meetings with employed workers are governed by a sequential-auction framework (Postel-Vinay and Robin 2002; Cahuc, Postel-Vinay, and Robin 2006), in which the incumbent and poaching firms compete for the worker.

Wages may also adjust in response to two types of stochastic shocks. First, following a productivity shock, the wage is adjusted only if the existing wage violates one party's participation constraint: it increases just enough to prevent the worker's post-shock value from falling below her unemployment value, or decreases just enough to prevent the firm's value from falling below the value of holding a vacancy. Otherwise, the wage remains unchanged (MacLeod and Malcomson 1993). Second, wages are renegotiated when learning opportunities are accepted, which occurs when the resulting increase in the worker's human capital raises the joint surplus. The parties then bargain over the incremental surplus.

The distributions of workers, firms, and matches evolve through exogenous shocks and endogenous

³Modeling cross-occupational meetings as a stochastic shock provides a tractable way to capture the imperfect transferability of skills across jobs. Workers who switch occupations cannot be expected to carry over all non-firm-specific skills acquired in previous employment, as some skills are effectively occupation-specific. The probability of a cross-occupational meeting and the associated haircut are disciplined by observed occupation switches and the wage changes of switchers (Section 4). This device therefore serves as a reduced-form proxy for occupation-specific human capital.

decisions. Workers and firms are subject to exogenous death shocks, while productivity shocks may destroy matches by reducing joint surplus below zero. Firms enter subject to a free-entry condition and exit when posting a vacancy is no longer profitable.

3.1. Workers, firms and matches

The model economy consists of two types of agents: a fixed unit measure of infinitely lived, risk-neutral workers and an endogenous measure of risk-neutral firms. Workers and firms can be matched in a one-to-one manner; when matched, the worker is employed and the firm is filled. Otherwise, the worker is unemployed and the firm is vacant. Unemployed workers engage in home production that yields a flow income, while vacant firms incur a vacancy flow cost. Both types of agents discount the future at rate r .

Each worker carries two stocks of human capital that grow while employed: portable human capital p and firm-specific human capital s . Both are continuously distributed across workers and observable by all agents. Portable human capital belongs to the worker: she retains it when a match ends, and it is productive at any employer. Firm-specific human capital is entirely lost when the match ends.⁴ Each worker is also endowed at entry with a permanent characteristic a , drawn from a distribution $\gamma_a(a)$: her *pre-market skill*, the skill she brings to the labor market from schooling and innate ability, before any human capital is accumulated on the job. It makes some workers more productive than others and shapes how fast they learn, in ways specified below. It is the only ex-ante difference between workers—everything else a worker carries, p and s , is accumulated on the job. Let $e(a, p, s)$ denote the density of type- (a, p, s) employed workers and $u(a, p)$ the density of type- (a, p) unemployed workers. The measures of employed and unemployed workers are $E = \iiint e(a, p, s) da dp ds$ and $U = \iint u(a, p) da dp$, with $E + U = 1$. Workers die exogenously at a rate δ_w .

Firms are defined by their labor productivity, z , and by the intensities with which they use portable and firm-specific human capital in production, ϕ_p and ϕ_s . Firm productivity and human capital usage intensities are continuously distributed across firms, and are observable by all agents. While firms' human capital usage intensities are fixed, their productivity is hit by idiosyncratic shocks that arrive at a rate δ_z . Upon the arrival of a shock, a new productivity level z' is drawn from a distribution with density γ_z , under which

$$\log z' \sim \mathcal{N}(0, \sigma_z^2),$$

where σ_z governs the dispersion of productivity draws. Between shocks productivity is constant, so its persistence is governed by how rarely shocks arrive: a firm keeps its current productivity for $1/\delta_z$ years on average.

⁴Formally, s is a state variable of the match rather than of the worker. Every new match starts with $s = 0$, the stock grows only within the match, and it has no counterpart once the match ends. Attaching it to the employed worker is a bookkeeping convention; accordingly, unemployed workers are described by (a, p) alone.

The density of type- (z, ϕ_p, ϕ_s) filled firms is $f(z, \phi_p, \phi_s)$, and the density of type- (z, ϕ_p, ϕ_s) vacant firms is $v(z, \phi_p, \phi_s)$. The measures of filled and vacant firms are $F = \iiint f(z, \phi_p, \phi_s) dz d\phi_p d\phi_s$ and $V = \iiint v(z, \phi_p, \phi_s) dz d\phi_p d\phi_s$ respectively, and the total measure of firms is $J = F + V$. Firms are destroyed exogenously at a rate δ_e , and, as explained below, may exit endogenously if the expected value of posting a vacancy is negative.

A match between a type- (a, p, s) worker and a type- (z, ϕ_p, ϕ_s) firm produces a flow of output

$$y(a, p, s, z, \phi_p, \phi_s) = z^\theta (a + \phi_p p + \phi_s s)^{1-\theta},$$

where $\theta \in (0, 1)$ governs the relative importance of firm productivity versus the worker's skill composite $a + \phi_p p + \phi_s s$ in production. Pre-market skill a enters the composite as a baseline stock, ensuring that output is strictly positive for any $a > 0$ even at the entry boundary $p = s = 0$; the usage intensities ϕ_p and ϕ_s weight the two accumulated stocks. Let $\Gamma(a, p, s, z, \phi_p, \phi_s)$ denote the equilibrium density of matches, such that

$$\iiint \Gamma(a, p, s, z, \phi_p, \phi_s) da dp ds = f(z, \phi_p, \phi_s), \quad (1)$$

and

$$\iiint \Gamma(a, p, s, z, \phi_p, \phi_s) dz d\phi_p d\phi_s = e(a, p, s). \quad (2)$$

3.2. Meeting technology and match formation

The number of matches between job seekers and vacancies is determined by a matching function $M : \mathbb{R}_+^2 \rightarrow \mathbb{R}_+$, which maps the total measure of job seekers, $U + \eta(1 - U)$, and the total measure of vacancies, V , into the number of meetings. Here, $\eta \in [0, 1]$ denotes the fraction of employed workers who are actively searching. The matching function is assumed to be Cobb–Douglas,

$$M(U + \eta(1 - U), V) = A(U + \eta(1 - U))^\mu V^{1-\mu},$$

where $A > 0$ measures matching efficiency and $\mu \in (0, 1)$ governs the elasticity of matches with respect to job seekers.

Let

$$\alpha = \frac{M(U + \eta(1 - U), V)}{[U + \eta(1 - U)]V} \quad (3)$$

denote the worker–vacancy meeting rate. The rate density at which an unmatched worker meets a type- (z, ϕ_p, ϕ_s) vacancy is $\alpha v(z, \phi_p, \phi_s)$, whereas the rate density at which a vacancy meets a type- (a, p) unmatched worker is $\alpha u(a, p)$. Similarly, matched workers meet type- (z, ϕ_p, ϕ_s) firms at rate density $\eta \alpha v(z, \phi_p, \phi_s)$, while the rate density at which vacancies meet type- (a, p, s) workers employed at a type- (z, ϕ_p, ϕ_s) firm is $\eta \alpha \Gamma(a, p, s, z, \phi_p, \phi_s)$.

Each meeting carries a label $\chi \in \{0, 1\}$ that flags the meeting as same-occupation ($\chi = 0$) or cross-occupation ($\chi = 1$). Labels are i.i.d. across meetings, with $\Pr(\chi = 1) = \pi_o \in [0, 1]$ and $\Pr(\chi = 0) = 1 - \pi_o$. On a same-occupation meeting, the match (if formed) starts with the worker's portable human capital p unchanged. On a cross-occupation meeting, the match (if formed) starts with portable human capital rescaled by a factor $\kappa \in (0, 1]$ at match formation. Define the effective portable human capital at the start of a candidate new match as

$$\tilde{p}(p, \chi) = \chi\kappa p + (1 - \chi)p. \quad (4)$$

Matches may be destroyed exogenously by worker death or firm destruction, with arrival rates δ_w and δ_e , respectively. They may also be endogenously destroyed if the worker is poached or if the surplus turns negative following a productivity shock. Workers cannot quit. Firms cannot replace a worker, nor can they reduce wages when a different worker would raise the firm's share of the surplus.

3.3. Wages

Wages vary only in response to stochastic events and remain fixed between such events. When an unemployed worker is hired, the initial wage grants her a share β of the match surplus. When an employed worker meets another vacant firm, the current employer and the poacher compete for her following the sequential-auction framework: she moves if the poaching match generates more value, and whichever match she does not take becomes the outside option that anchors her wage. When firm productivity changes, the wage adjusts only if forced—raised to keep the worker in the match if the old wage leaves her worse off than unemployment, or lowered to keep the firm in the match if the old wage grants the worker more than the entire new surplus. And when a learning opportunity is accepted, the parties bargain over the surplus it creates. The remainder of this subsection defines the relevant objects and derives the wage in each case.

3.3.1. Match surplus

The joint match surplus is the key object in wage determination. It measures the total value created by a match relative to the outside options of the two parties. Let $W_u(a, p)$ denote the present value of unemployment for a type- (a, p) unmatched worker, and $\Pi_v(z, \phi_p, \phi_s)$ the present value of posting a vacancy. Let $W_e(a, p, s, z, \phi_p, \phi_s, w)$ and $\Pi_f(a, p, s, z, \phi_p, \phi_s, w)$ denote, respectively, the present values of a wage contract w for a type- (a, p, s) worker and a type- (z, ϕ_p, ϕ_s) firm that are matched to each other. For a given match between a type- (a, p, s) worker and a type- (z, ϕ_p, ϕ_s) firm, the joint match

surplus is

$$S(a, p, s, z, \phi_p, \phi_s) = \Pi_f(a, p, s, z, \phi_p, \phi_s, w) - \max\{\Pi_v(z, \phi_p, \phi_s), 0\} \\ + W_e(a, p, s, z, \phi_p, \phi_s, w) - W_u(a, p). \quad (5)$$

Although the values of the matched worker and the matched firm each depend on the wage, their sum—the joint match surplus—does not, because the wage is a pure transfer between the two parties and endogenous decisions do not depend on it. The firm's outside option upon match dissolution is $\max\{\Pi_v, 0\}$ rather than Π_v , reflecting the firm's option to exit when posting at (z, ϕ_p, ϕ_s) would be unprofitable.

To simplify the exposition, let

$$L_u(a, p) \equiv W_u(a, p) - W_u(a, \kappa p), \quad (6)$$

and define the *effective surplus* $\tilde{S}$ for a worker of type (a, p) meeting a firm (z, ϕ_p, ϕ_s) with meeting label $\chi \in \{0, 1\}$ as

$$\tilde{S}(a, p, s, z, \phi_p, \phi_s; \chi) = \begin{cases} S(a, p, s, z, \phi_p, \phi_s) & \text{if } \chi = 0, \\ S(a, \kappa p, s, z, \phi_p, \phi_s) - L_u(a, p) & \text{if } \chi = 1. \end{cases} \quad (7)$$

The effective surplus $\tilde{S}$ is the quantity on which matching decisions are based: a meeting with an unemployed worker produces a match only if $\tilde{S}$ is positive, and an employed worker moves to the firm she has met only if the new match's $\tilde{S}$ exceeds her current match's surplus. On a same-occupation meeting, it is simply the match surplus. On a cross-occupation meeting, the surplus of the prospective match is evaluated at the reduced stock κp , while the worker's outside option remains her value of unemployment at the full stock p ; subtracting L_u accounts for this asymmetry, so a cross-occupation match is worth forming only if the surplus it creates exceeds what the worker gives up through the haircut. Writing decisions in terms of $\tilde{S}$ lets every expression that follows treat the two meeting labels symmetrically.

3.3.2. Unemployed workers

Let $w_0(a, p, z, \phi_p, \phi_s, \chi)$ denote the initial wage of a worker when transitioning from unemployment at (a, p) to a match with a type- (z, ϕ_p, ϕ_s) firm. The wage is determined by Nash bargaining and splits the effective match surplus according to the worker's bargaining power β ,

$$W_e(a, \tilde{p}, 0, z, \phi_p, \phi_s, w_0) - W_u(a, p) = \beta \tilde{S}(a, p, 0, z, \phi_p, \phi_s; \chi). \quad (8)$$

3.3.3. Poaching

A poaching event occurs when a type- (a, p, s) worker matched with a type- (z, ϕ_p, ϕ_s) firm meets a vacancy at (z', ϕ'_p, ϕ'_s) with label $\chi \in \{0, 1\}$, leading to one out of three possible outcomes.

First, if the effective surplus at the poaching firm is lower than the worker's current share of the surplus at the current employer, i.e., if $\tilde{S}(a, p, 0, z', \phi'_p, \phi'_s; \chi) < W_e(a, p, s, z, \phi_p, \phi_s, w) - W_u(a, p) \leq S(a, p, s, z, \phi_p, \phi_s)$, the worker cannot credibly threaten to leave. No renegotiation occurs and the worker remains at the current employer with the same wage.

Second, if the effective surplus at the poaching firm exceeds the worker's current share of the match surplus but is weakly less than the surplus of the current match, i.e.,

$$W_e(a, p, s, z, \phi_p, \phi_s, w) - W_u(a, p) < \tilde{S}(a, p, 0, z', \phi'_p, \phi'_s; \chi) \leq S(a, p, s, z, \phi_p, \phi_s),$$

the current employer retains the worker and the wage is renegotiated upward. The bargained wage w_1^X satisfies

$$\begin{aligned} W_e(a, p, s, z, \phi_p, \phi_s, w_1^X) - W_u(a, p) &= \tilde{S}(a, p, 0, z', \phi'_p, \phi'_s; \chi) \\ &\quad + \beta [S(a, p, s, z, \phi_p, \phi_s) - \tilde{S}(a, p, 0, z', \phi'_p, \phi'_s; \chi)] \\ &= \beta S(a, p, s, z, \phi_p, \phi_s) + (1 - \beta) \tilde{S}(a, p, 0, z', \phi'_p, \phi'_s; \chi). \end{aligned} \quad (9)$$

Lastly, if the effective surplus at the poaching firm exceeds the surplus of the current match, the worker switches jobs. The bargained wage w_2^X at the new match satisfies

$$\begin{aligned} W_e(a, \tilde{p}, 0, z', \phi'_p, \phi'_s, w_2^X) - W_u(a, p) &= S(a, p, s, z, \phi_p, \phi_s) \\ &\quad + \beta [\tilde{S}(a, p, 0, z', \phi'_p, \phi'_s; \chi) - S(a, p, s, z, \phi_p, \phi_s)] \\ &= \beta \tilde{S}(a, p, 0, z', \phi'_p, \phi'_s; \chi) + (1 - \beta) S(a, p, s, z, \phi_p, \phi_s). \end{aligned} \quad (10)$$

3.3.4. Productivity shocks

Firms are subject to idiosyncratic productivity shocks with arrival rate δ_z . When a shock occurs, a new productivity level z' is drawn from $\gamma_z(z')$. The match then produces output $y(a, p, s, z', \phi_p, \phi_s)$, and the joint surplus is $S(a, p, s, z', \phi_p, \phi_s)$.

If the new surplus is negative or the firm's continuation value as a vacancy at the new productivity is negative, i.e., $S(a, p, s, z', \phi_p, \phi_s) < 0$ or $\Pi_v(z', \phi_p, \phi_s) < 0$, the match is destroyed and the worker becomes unemployed, losing accumulated firm-specific human capital. The firm turns vacant if $\Pi_v(z', \phi_p, \phi_s) \geq 0$ and exits otherwise. Exit is a decision of the firm as a whole: when a productivity shock makes the firm unprofitable even as a vacancy, the firm closes and its match ends with it,

which ties firm closures to productivity and gives the model endogenous closures alongside exogenous destruction.

If the match survives—that is, if $S(a, p, s, z', \phi_p, \phi_s) \geq 0$ and $\Pi_v(z', \phi_p, \phi_s) \geq 0$ —it continues to produce, but the wage contract may need to be renegotiated.

If $W_e(a, p, s, z', \phi_p, \phi_s, w) - W_u(a, p) < 0$, a new wage w^+ is set such that

$$W_e(a, p, s, z', \phi_p, \phi_s, w^+) - W_u(a, p) = 0. \quad (11)$$

That is, if the pre-shock wage leaves the worker worse off than unemployment, the wage is raised to make the worker indifferent.

Alternatively, if $W_e(a, p, s, z', \phi_p, \phi_s, w) - W_u(a, p) > S(a, p, s, z', \phi_p, \phi_s)$, a new wage w^- is set such that

$$W_e(a, p, s, z', \phi_p, \phi_s, w^-) - W_u(a, p) = S(a, p, s, z', \phi_p, \phi_s). \quad (12)$$

In other words, if the pre-shock wage grants the worker more than the entire surplus, the wage is reduced so that the worker captures exactly the surplus.

Lastly, if $0 \leq W_e(a, p, s, z', \phi_p, \phi_s, w) - W_u(a, p) \leq S(a, p, s, z', \phi_p, \phi_s)$, the wage remains unchanged.

3.3.5. Learning opportunities

Matches receive learning opportunities at rate δ_l . A learning opportunity may increase the employed worker's human capital, thereby affecting both the match's output and the worker's outside option. An opportunity arrives carrying a human capital increment $h \geq 0$, drawn from an exponential distribution,

$$\begin{aligned} \gamma_h(h \mid a, p, s, \phi_p, \phi_s) &= \lambda_h(a, p, s, \phi_p, \phi_s) \exp(-\lambda_h(a, p, s, \phi_p, \phi_s)h), \\ \lambda_h(a, p, s, \phi_p, \phi_s) &= \frac{1}{(a + \phi_p p + \phi_s s)^\ell}, \end{aligned}$$

so that the expected increment depends on the worker's pre-market skill, her accumulated stocks, and the employing firm's usage intensities, with curvature governed by ℓ ; pre-market skill a enters the hazard additively, keeping it finite and bounded away from zero even at the entry boundary $p = s = 0$. Observing h , the worker and the firm jointly decide whether to undertake the opportunity; it is taken only if both parties agree.

If the opportunity is undertaken, the increment is allocated between portable and firm-specific human capital according to the firm's human capital usage intensities:

$$p' = p + \Delta(\phi_p, \phi_s)h, \quad s' = s + (1 - \Delta(\phi_p, \phi_s))h, \quad \Delta(\phi_p, \phi_s) = \frac{\phi_p}{\phi_p + \phi_s}.$$

Firms that use firm-specific human capital more intensively therefore tilt learning toward it, and portable-intensive firms toward portable human capital. Acceptance triggers a renegotiation of the wage.⁵ The worker and the firm bargain over the incremental surplus generated by the additional human capital, taking the current contract as the disagreement point. The worker retains the surplus she enjoyed before the learning opportunity—measured relative to her contemporaneous outside option—and receives an additional fraction β of the increase in joint surplus. The bargained wage w_3 therefore satisfies

$$W_e(a, p', s', z, \phi_p, \phi_s, w_3) - W_u(a, p') = W_e(a, p, s, z, \phi_p, \phi_s, w) - W_u(a, p) + \beta(S(a, p', s', z, \phi_p, \phi_s) - S(a, p, s, z, \phi_p, \phi_s)). \quad (13)$$

This bargaining rule implies that the worker's gain from accepting the opportunity is

$$\begin{aligned} & W_e(a, p', s', z, \phi_p, \phi_s, w_3) - W_e(a, p, s, z, \phi_p, \phi_s, w) \\ &= [W_u(a, p') - W_u(a, p)] + \beta(S(a, p', s', z, \phi_p, \phi_s) - S(a, p, s, z, \phi_p, \phi_s)), \end{aligned}$$

while the firm's gain is

$$\begin{aligned} & \Pi_f(a, p', s', z, \phi_p, \phi_s, w_3) - \Pi_f(a, p, s, z, \phi_p, \phi_s, w) \\ &= (1 - \beta)(S(a, p', s', z, \phi_p, \phi_s) - S(a, p, s, z, \phi_p, \phi_s)). \end{aligned}$$

The decomposition of the worker's gain is informative. She receives two components. First, she obtains her bargained share of the increase in match surplus. Second, she captures the entire improvement in her outside option, $W_u(a, p') - W_u(a, p)$, because the retained surplus is evaluated relative to the outside option after the new human capital has been acquired. By contrast, the firm's gain is proportional only to the change in joint surplus.

These expressions immediately characterize the acceptance decision. Since the worker's gain includes the outside-option improvement, she is willing to accept every opportunity that raises joint surplus and, in addition, some opportunities that reduce it. The firm, however, benefits only when joint surplus weakly increases. Consequently, mutual consent occurs if and only if

$$S(a, p', s', z, \phi_p, \phi_s) \geq S(a, p, s, z, \phi_p, \phi_s),$$

so the equilibrium acceptance rule is simply to undertake learning whenever it weakly raises the match

⁵Learning triggers a renegotiation while a productivity shock renegotiates only when forced (Section 3.3.4) because the two events differ in nature: a productivity shock is imposed, requiring no one's consent, whereas learning happens only if both parties agree, and a consent decision is an occasion to bargain. Once consent is required, renegotiation is also what preserves the wage independence of decisions: were the wage held fixed at w , each party's consent would depend on w , endogenous decisions would no longer depend on the state alone, and the joint surplus would cease to be independent of the wage.

surplus. In equilibrium, therefore, every rejection is driven by the firm. Rejection occurs when the increase in the worker's outside option exceeds the increase in the pair's joint value, so that learning raises the worker's value but lowers the surplus available to the match.

3.4. Value functions

The equilibrium of the model is characterized by a set of value functions describing the expected discounted payoffs of workers and firms in each possible state. These value functions incorporate current payoffs, the arrival of shocks, and endogenous transitions due to match formation, renegotiation, and separation.

Workers' value functions distinguish between unemployment and employment. The value of unemployment reflects home production and the expected gains from meeting vacancies, while the value of employment depends on the wage, the characteristics of the firm, the worker's accumulated human capital, and the possibility of poaching, productivity shocks, learning opportunities, and match destruction.

Firms' value functions distinguish between vacant and filled positions. The value of a vacancy reflects posting costs, meeting probabilities, and the option to exit, while the value of a filled firm depends on current production, wages, and the expected evolution of the match.

The match surplus governs match formation, continuation, destruction, and wage determination. To compress notation, define the label-averaged effective surplus operator, for any constant $x \in \mathbb{R}$,

$$\begin{aligned} \Psi(a, p, z, \phi_p, \phi_s; x) := & (1 - \pi_o) \max \{ \tilde{S}(a, p, 0, z, \phi_p, \phi_s; 0) - x, 0 \} \\ & + \pi_o \max \{ \tilde{S}(a, p, 0, z, \phi_p, \phi_s; 1) - x, 0 \}, \end{aligned} \quad (14)$$

which is the expected effective surplus, net of an outside-option benchmark x , of a meeting with worker state (a, p) and vacancy (z, ϕ_p, ϕ_s) , taking expectations over the meeting label.

3.4.1. Unemployed workers

The flow value of unemployment for a type- (a, p) unmatched worker satisfies

$$\begin{aligned} (r + \delta_w) W_u(a, p) \\ = b + \alpha \beta \iiint \Psi(a, p, z', \phi'_p, \phi'_s; 0) v(z', \phi'_p, \phi'_s) dz' d\phi'_p d\phi'_s, \end{aligned} \quad (15)$$

where b is the flow income from home production. The second term is the worker's expected share of the effective surplus over the joint distribution of vacancy types and meeting labels.

3.4.2. Vacant firms

The flow value of a vacancy for a type- (z, ϕ_p, ϕ_s) firm satisfies

$$\begin{aligned}
(r + \delta_e)\Pi_v(z, \phi_p, \phi_s) = & -c_v + \delta_z \int \left(\max\{\Pi_v(z', \phi_p, \phi_s), 0\} - \Pi_v(z, \phi_p, \phi_s) \right) \gamma_z(z') dz' \\
& + \alpha(1 - \beta) \iint \Psi(a, p, z, \phi_p, \phi_s; 0) u(a, p) da dp \\
& + \eta\alpha(1 - \beta) \int \dots \int \Psi(a, p, z, \phi_p, \phi_s; S(a, p, s, z', \phi'_p, \phi'_s)) \\
& \times \Gamma(a, p, s, z', \phi'_p, \phi'_s) da dp ds dz' d\phi'_p d\phi'_s,
\end{aligned} \tag{16}$$

where c_v is the flow cost of keeping a vacancy open. The second term captures the effect of productivity transitions from z to z' , with the $\max\{\cdot, 0\}$ reflecting the vacancy's option to exit when the new productivity makes vacancy posting unprofitable. The third term represents the vacancy's expected payoff from matching with an unmatched worker, integrated over the worker's (a, p) and the meeting label. The fourth term reflects the expected gain from poaching an employed worker, where the relevant comparison surplus is the surplus at the worker's current match $S(a, p, s, z', \phi'_p, \phi'_s)$.

3.4.3. Employed workers, filled firms and joint surplus

Appendix B.1 shows that the net flow value of employment for a type- (a, p, s) worker matched with a type- (z, ϕ_p, ϕ_s) firm under a wage contract w satisfies

$$\begin{aligned}
& (r + \delta_w + \delta_e) [W_e(a, p, s, z, \phi_p, \phi_s, w) - W_u(a, p)] \\
& = w - b - \alpha\beta \iiint \Psi(a, p, z', \phi'_p, \phi'_s; 0) v(z', \phi'_p, \phi'_s) dz' d\phi'_p d\phi'_s \\
& + \eta\alpha \Xi_W(a, p, s, z, \phi_p, \phi_s, w) \\
& + \delta_z \left[\int_{\mathcal{Z}_0} [W_u(a, p) - W_e(a, p, s, z, \phi_p, \phi_s, w)] \gamma_z(z') dz' \right. \\
& + \int_{\mathcal{Z}_1} [W_e(a, p, s, z', \phi_p, \phi_s, w^+) - W_e(a, p, s, z, \phi_p, \phi_s, w)] \gamma_z(z') dz' \\
& + \int_{\mathcal{Z}_2} [W_e(a, p, s, z', \phi_p, \phi_s, w^-) - W_e(a, p, s, z, \phi_p, \phi_s, w)] \gamma_z(z') dz' \\
& \left. + \int_{\mathcal{Z}_3} [W_e(a, p, s, z', \phi_p, \phi_s, w) - W_e(a, p, s, z, \phi_p, \phi_s, w)] \gamma_z(z') dz' \right] \\
& + \delta_l \int_{\mathcal{H}} [W_e(a, p + \Delta(\phi_p, \phi_s)h, s + (1 - \Delta(\phi_p, \phi_s))h, z, \phi_p, \phi_s, w_3)
\end{aligned}$$

$$- W_e(a, p, s, z, \phi_p, \phi_s, w)] \gamma_h(h|a, p, s, \phi_p, \phi_s) dh, \quad (17)$$

where the poaching term is

$$\begin{aligned} \Xi_W(a, p, s, z, \phi_p, \phi_s, w) = & (1 - \pi_o) \iiint_{\mathcal{P}_1(0)} [W_e(a, p, s, z, \phi_p, \phi_s, w_1^0) \\ & - W_e(a, p, s, z, \phi_p, \phi_s, w)] \nu dz' d\phi'_p d\phi'_s \\ & + (1 - \pi_o) \iiint_{\mathcal{P}_2(0)} [W_e(a, p, 0, z', \phi'_p, \phi'_s, w_2^0) \\ & - W_e(a, p, s, z, \phi_p, \phi_s, w)] \nu dz' d\phi'_p d\phi'_s \\ & + \pi_o \iiint_{\mathcal{P}_1(1)} [W_e(a, p, s, z, \phi_p, \phi_s, w_1^1) \\ & - W_e(a, p, s, z, \phi_p, \phi_s, w)] \nu dz' d\phi'_p d\phi'_s \\ & + \pi_o \iiint_{\mathcal{P}_2(1)} [W_e(a, \kappa p, 0, z', \phi'_p, \phi'_s, w_2^1) \\ & - W_e(a, p, s, z, \phi_p, \phi_s, w)] \nu dz' d\phi'_p d\phi'_s. \end{aligned} \quad (18)$$

The integration regions are

$$\begin{aligned} \mathcal{P}_1(\chi) = & \{(z', \phi'_p, \phi'_s) : W_e(a, p, s, z, \phi_p, \phi_s, w) - W_u(a, p) \\ & < \tilde{S}(a, p, 0, z', \phi'_p, \phi'_s; \chi) \leq S(a, p, s, z, \phi_p, \phi_s)\}, \\ \mathcal{P}_2(\chi) = & \{(z', \phi'_p, \phi'_s) : \tilde{S}(a, p, 0, z', \phi'_p, \phi'_s; \chi) > S(a, p, s, z, \phi_p, \phi_s)\}, \end{aligned}$$

where $\tilde{S}$ is defined in (7) and w_1^χ, w_2^χ are the bargained wages from (9)–(10), and

$$\begin{aligned} \mathcal{Z}_0 = & \{z' : S(a, p, s, z', \phi_p, \phi_s) < 0 \text{ or } \Pi_\nu(z', \phi_p, \phi_s) < 0\}, \\ \mathcal{Z}_1 = & \{z' \notin \mathcal{Z}_0 : W_e(a, p, s, z', \phi_p, \phi_s, w) - W_u(a, p) < 0\}, \\ \mathcal{Z}_2 = & \{z' \notin \mathcal{Z}_0 : W_e(a, p, s, z', \phi_p, \phi_s, w) - W_u(a, p) > S(a, p, s, z', \phi_p, \phi_s)\}, \\ \mathcal{Z}_3 = & \{z' \notin \mathcal{Z}_0 : 0 \leq W_e(a, p, s, z', \phi_p, \phi_s, w) - W_u(a, p) \leq S(a, p, s, z', \phi_p, \phi_s)\}, \\ \mathcal{H} = & \{h : S(a, p + \Delta h, s + (1 - \Delta)h, z, \phi_p, \phi_s) \geq S(a, p, s, z, \phi_p, \phi_s)\}, \end{aligned}$$

with $\Delta = \Delta(\phi_p, \phi_s)$ and w^+, w^-, w_3 defined by (11)–(12), (13). The region $\mathcal{Z}_0$ collects both destruction triggers of Section 3.3.4—a post-shock surplus that turns negative and a post-shock vacancy value that turns negative, the latter destroying the match even when the surplus would remain positive—while $\mathcal{Z}_1$ – $\mathcal{Z}_3$ partition the surviving states according to whether the wage must be raised, cut, or left unchanged.

The filled firm's net flow value is the mirror image of the worker's: the same events drive it, the wage enters with the opposite sign, and the firm's fallback is the vacancy value—with the option to exit—in place of unemployment. Appendix B.2 reports it in full as equation (A3).

Lastly, Appendix B.3 shows that equations (17) and (A3) can be combined to express the surplus $S(a, p, s, z, \phi_p, \phi_s)$ as the fixed point of

$$\begin{aligned}
& (r + \delta_w + \delta_e)S(a, p, s, z, \phi_p, \phi_s) \\
&= y(a, p, s, z, \phi_p, \phi_s) - b + c_v \\
&- \alpha(1 - \beta) \iint \Psi(a', p', z, \phi_p, \phi_s; 0) u(a', p') da' dp' \\
&- \alpha\beta \iiint \Psi(a, p, z', \phi'_p, \phi'_s; 0) v(z', \phi'_p, \phi'_s) dz' d\phi'_p d\phi'_s \\
&- \eta\alpha(1 - \beta) \int \dots \int \Psi(a', p', z, \phi_p, \phi_s; S(a', p', s', z', \phi'_p, \phi'_s)) \\
&\quad \times \Gamma(a', p', s', z', \phi'_p, \phi'_s) da' dp' ds' dz' d\phi'_p d\phi'_s \\
&+ \eta\alpha\beta \iiint \Psi(a, p, z', \phi'_p, \phi'_s; S(a, p, s, z, \phi_p, \phi_s)) v(z', \phi'_p, \phi'_s) dz' d\phi'_p d\phi'_s \\
&+ \delta_z \int \left[\mathbf{1}\{\Pi_v(z', \phi_p, \phi_s) \geq 0\} \max\{S(a, p, s, z', \phi_p, \phi_s), 0\} \right. \\
&\quad \left. - S(a, p, s, z, \phi_p, \phi_s) \right] \gamma_z(z') dz' \\
&+ \delta_l \int \mathbf{1}\{S(a, p + \Delta h, s + (1 - \Delta)h, z, \phi_p, \phi_s) \geq S(a, p, s, z, \phi_p, \phi_s)\} \\
&\quad \times (S(a, p + \Delta h, s + (1 - \Delta)h, z, \phi_p, \phi_s) - S(a, p, s, z, \phi_p, \phi_s)) \\
&\quad + W_u(a, p + \Delta h) - W_u(a, p) \gamma_h(h|a, p, s, \phi_p, \phi_s) dh, \tag{19}
\end{aligned}$$

with $\Delta = \Delta(\phi_p, \phi_s)$.

3.5. Entry and exit

Workers may exit exogenously upon hitting a death shock with arrival rate δ_w . Exiting workers are replaced by unemployed entrants who have no human capital of any type and draw their pre-market skill a from a Beta distribution on $[0, 1]$ with shape parameters $a_\alpha, a_\beta > 0$, whose density is denoted $\gamma_a(a)$.⁶ Workers do not exit endogenously.

Firms may exit exogenously if hit with a firm destruction shock at arrival rate δ_e . Additionally, a firm may exit endogenously if the expected value of posting a vacancy is negative. Exiting firms are replaced by entrants. An entrant draws its productivity from γ_z , and its human capital usage intensities jointly

⁶I normalize the scale of pre-market skill by setting $a \in [0, 1]$. Because every targeted moment is invariant to the scale of output, only relative variation in a is identified, not its absolute scale.

from a bivariate log-normal distribution,

$$\begin{pmatrix} \log \phi_p \\ \log \phi_s \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_p \\ \mu_s \end{pmatrix}, \begin{pmatrix} \sigma_p^2 & \rho_\phi \sigma_p \sigma_s \\ \rho_\phi \sigma_p \sigma_s & \sigma_s^2 \end{pmatrix}\right),$$

whose joint density is denoted $\gamma_\phi(\phi_p, \phi_s)$. The measure of firms J is determined through a free-entry condition whereby potential entrants face an entry cost c_e . Upon learning their type, firms decide whether to exit immediately or to post a vacancy. In an equilibrium with a positive measure of firms,

$$\iiint \max\{\Pi_v(z, \phi_p, \phi_s), 0\} \gamma_z(z) \gamma_\phi(\phi_p, \phi_s) dz d\phi_p d\phi_s \leq c_e \quad (20)$$

holds with equality.

3.6. Steady-state equilibrium

In a stationary equilibrium, $u(a, p)$, $v(z, \phi_p, \phi_s)$, and $\Gamma(a, p, s, z, \phi_p, \phi_s)$ —and consequently $e(a, p, s)$ and $f(z, \phi_p, \phi_s)$ —are invariant.

Stationarity requires that, for every state, inflows balance outflows. The balance equations—for the match density Γ , including the boundary condition at $s = 0$, where every new match enters; for the unemployed density u ; for the vacancy density v ; and for the entry rate μ_J that keeps the measure of firms constant—are stated and discussed in Appendix C.

3.6.1. Equilibrium definition

A stationary equilibrium consists of value functions W_u, W_e, Π_v, Π_f and surplus S satisfying (15)–(19); stationary densities u, v, Γ satisfying (A5)–(A8), with v supported on $\{(z, \phi_p, \phi_s) : \Pi_v(z, \phi_p, \phi_s) \geq 0\}$ and Γ supported on $\{(a, p, s, z, \phi_p, \phi_s) : S(a, p, s, z, \phi_p, \phi_s) \geq 0 \text{ and } \Pi_v(z, \phi_p, \phi_s) \geq 0\}$; a total measure of firms J satisfying the free-entry condition (20) holding with equality; an entry rate μ_J satisfying the balance condition (A9); and a meeting rate α consistent with (3) given u, v, J .

4. Estimation

This section takes the model to the data. I estimate the structural parameters using the simulated method of moments (SMM). The parameters solve

$$\hat{\vartheta} = \arg \min_{\vartheta} [m_{\text{data}} - m_{\text{sim}}(\vartheta)]' W [m_{\text{data}} - m_{\text{sim}}(\vartheta)], \quad (21)$$

where m_{data} is the vector of empirical moments and $m_{\text{sim}}(\vartheta)$ is its model-simulated counterpart, obtained by solving the model at ϑ , simulating a panel with the same spell-based, annual structure as the data, and applying to it the same moment calculations used on the data. The weighting matrix W is diagonal: each moment's weight is the inverse of the sum of its sampling variance in the data, estimated by bootstrap, and the simulation variance of its model counterpart, so that cells that are noisy on either side are automatically downweighted.

Two features of the empirical panel shape the comparison between the data and the model-generated panels. First, wages are observed once per calendar year within each spell: a match that spans several years generates one observation per year, and its wage changes are observed only as the net change over the calendar year. The simulated panel is aggregated to the same spell-year structure before any moment is computed, so this attenuation is identical on both sides; it reduces the precision of the wage-based moments but does not bias the estimates. Second, wages and output per worker enter every worker-panel moment in the adjusted form of Section 2.5: for wages, variation associated with gender, region, industry, calendar year and occupation is removed; for output per worker, variation associated with industry and calendar year. Each residual is then re-centered at its sample mean, and the worker's share of output is computed from the two. The model has no counterparts to gender, geography, industry or calendar year, and it treats occupations symmetrically, so it generates no systematic differences in pay across them. The residualization is therefore applied to the data only, and leaves the variation that the model is meant to explain.

Lastly, reported hourly wages are measured with error in the data. Therefore, before computing any statistic on the simulated panel, an i.i.d. normal error term with standard deviation σ_{me} is added to each simulated log wage. The parameter σ_{me} is estimated jointly with the model parameters and is identified by the lag-1 autocorrelation of within-match wage changes.

4.1. Parameters

The model is characterized by the following structural parameters. Workers and firms meet through a matching function with efficiency A and elasticity μ , and employed workers search with intensity η relative to the unemployed. A meeting crosses occupations with probability π_o , in which case the worker keeps a fraction κ of her portable human capital. Wages are set with worker bargaining power β , and an unemployed worker produces b at home. Matches end when the worker leaves the labor market, at rate δ_w , or the firm is destroyed, at rate δ_e . Firm productivity is redrawn at rate δ_z from a distribution with dispersion σ_z and enters output with weight θ , and the worker's skill composite with weight $1 - \theta$. Learning opportunities arrive at rate δ_l , with increments whose curvature in the accumulated stock is ℓ . Entering workers draw their pre-market skill from a Beta distribution with shape parameters (a_α, a_β) , and entering firms draw their log usage intensities $(\log \phi_p, \log \phi_s)$ from a bivariate normal distribution with means μ_p and μ_s , standard deviations σ_p and σ_s , and correlation

ρ_ϕ . Posting a vacancy costs c_v per unit of time, entering the market costs c_e , and all agents discount the future at rate r .

The discount rate r is set to 0.05 by convention, following the search-and-matching literature. The matching elasticity μ is not separately identified in a stationary equilibrium with constant aggregate unemployment and vacancies; it is set to one half following Petrongolo and Pissarides (2001). The mean of log productivity is normalized to zero, fixing the scale of z ; because every targeted moment is unchanged when the scale of output changes, this location is not separately identified. The remaining 22 parameters are estimated jointly by SMM.

4.2. Identification

The parameters are estimated jointly, so most moments respond to several parameters. Nevertheless, it is useful to provide a heuristic account of identification by pairing each parameter with its most informative moment, as shown in Table 2. I focus on three blocks of parameters that govern wage determination in the model: the productivity process, the learning technology, and the distribution of usage intensities across jobs. Appendix E provides the corresponding discussion for the remaining parameters.

Productivity dynamics from firm-level output. Productivity dynamics are inferred from changes in value added per worker within firms. In the model, productivity is a match-level attribute; in the data, value added is measured at the firm level, averaged across all positions. Interpreting the latter as a measure of the former requires two assumptions.

The first is that changes in firm productivity affect the value added of all positions proportionally. The forces behind firm-level changes—a new market, a change in product prices, a new technology or investment, or a change in management—operate at the firm level. Positions may be exposed to these forces differently, but productivity in the model captures their common component. If differences in exposure are unrelated to the human capital of the worker in each position, the firm-wide average captures the change that the model attributes to each match.

The second assumption is that, net of industry-year effects, year-to-year changes in firm value added per worker reflect only productivity redraws, not changes in workforce composition or movements in value added unrelated to productivity. Three features of the measurement support this assumption. First, the moments are computed across workers, giving greater weight to large firms, whose average workforce changes little from one year to the next. Second, residualizing on industry and year removes sectoral cycles and price movements; trimming changes at the 1st and 99th percentiles limits the influence of extreme accounting fluctuations; and within-firm differences remove permanent cross-firm differences, such as capital intensity. Third, an investment that raises value added per worker constitutes a productivity change in the model, which measures output per unit of worker composite,

rather than total factor productivity.

Under both assumptions, the within-firm change in value added per worker is the change in productivity plus the learning of the firm's workers, which is small and, as the next paragraph explains, pinned down by wages. If the second assumption fails, workforce turnover and accounting noise introduce small, frequent changes into measured productivity. The resulting estimates overstate the frequency of productivity redraws and understate their size.

The within-match autocorrelation of log value added per worker identifies θ . A shock resets the firm's productivity, so the persistence of output between shocks comes from the component they leave unchanged: the worker's composite, which evolves slowly through learning. A higher θ means productivity accounts for a larger share of output, so each shock renews a larger fraction of output and reduces its persistence.

The distribution of annual changes in value added per worker identifies δ_z . Shocks arrive discretely. When shocks are rare, most matches experience little change, while a few experience large jumps. This produces a sharply peaked distribution with long tails. Conditional on the variance of annual changes, the spread between the 10th and 90th percentiles therefore identifies the frequency of shocks. Finally, the variance of annual changes identifies σ_z . Given the frequency of shocks and the share of output renewed by each shock, the variance reveals the magnitude of productivity changes.

Figure 1 shows this variation. Each panel plots a statistic at the estimated parameters and after moving one parameter, the others held at their estimates. More frequent productivity shocks, with their dispersion lowered to keep the variance of annual changes fixed, flatten the peak of the distribution of changes and thicken its shoulders, which widens the gap between its 90th and 10th percentiles (Panel A). A larger weight on productivity lowers the autocorrelation of output at every lag (Panel B).

Learning from within-match wage growth. The learning technology is identified from wage dynamics within continuing matches, after accounting for the two other sources of within-match wage growth. The first is firm productivity. Wages are largely insulated from productivity shocks because these shocks trigger wage adjustments only when necessary (Section 3.3.4): wages rise if the existing wage leaves the worker worse off than unemployment and fall if the current wage would make the worker's share exceed the new match surplus. Such adjustments are rare, since a shock changes the wage only when it violates one party's participation constraint, and they contribute little to the mean of within-match wage growth.

The second source is the job ladder: an outside offer can raise a worker's wage even if she remains with her current employer (Section 3.3). The frequency and magnitude of these raises depend on parameters identified by other moments. The frequency with which an employed worker meets a firm depends on A and η , which are identified by job-finding and job-to-job transition rates. The probability that an outside offer generates a raise depends on the distribution of surplus across firms, which is

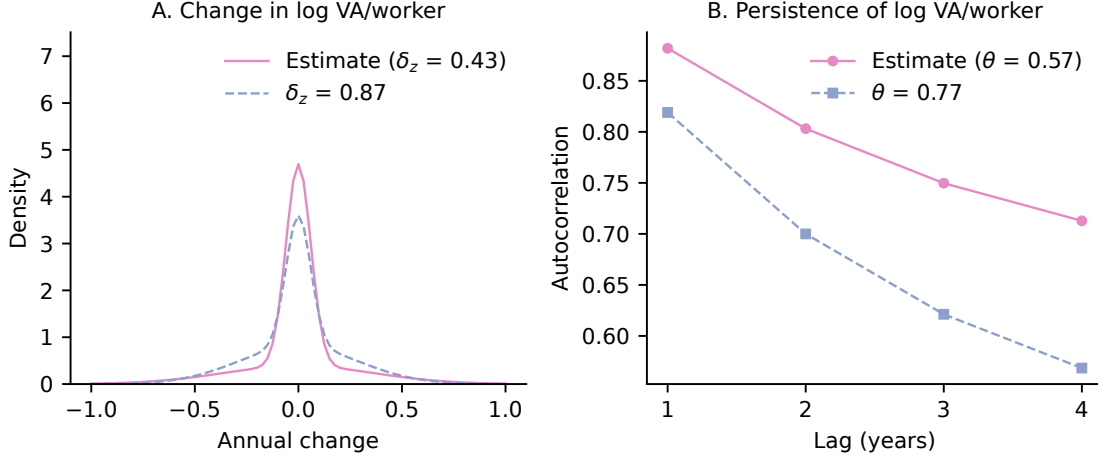


FIGURE 1. Identifying variation: the productivity process.

NOTES: Each statistic is computed on panels simulated from the model at the estimated parameters (solid line) and after moving one parameter, the others held at their estimates (dashed line); the legends give the value of the parameter moved. Panel A: density of annual changes in log value added per worker within matches; when δ_z moves, σ_z moves with it to hold the variance of the changes at its estimated value. Panel B: autocorrelation of log value added per worker within matches, by lag in years.

determined by the productivity process and the distribution of usage intensities. The share of the gain captured by the worker depends on β , which is identified by the worker's mean share of output. Given these parameters, the model determines the contribution of the job ladder to within-match wage growth, and the residual identifies the contribution of learning.

The mean of this residual identifies δ_l : because learning arrives in discrete increments, more frequent learning raises mean wage growth within a match. Wage growth among workers with ten or more years of tenure then identifies the curvature parameter ℓ . Because the expected learning increment depends on the accumulated human capital stock, the extent to which wage growth at long tenure falls below or rises above its average reveals whether workers with greater accumulated human capital learn more slowly or more quickly than beginners.

Both learning-driven wage growth and job-ladder raises decline with tenure, but for different reasons: the former reflects changes in learning increments as human capital accumulates, while the latter reflects the declining likelihood that workers receive outside offers worth matching as they climb the job ladder. The two are therefore told apart by the flow moments that fix the size of the ladder, not by the tenure profile alone; the untargeted bins in Figure 5 then check the shape of the wage-growth profile implied by this decomposition.

Figure 2 shows how the two learning parameters move the tenure profile of within-match wage growth, the others held at their estimates. More frequent learning opportunities raise wage growth at every tenure (Panel A), while a less negative curvature raises it mostly at long tenure (Panel B).

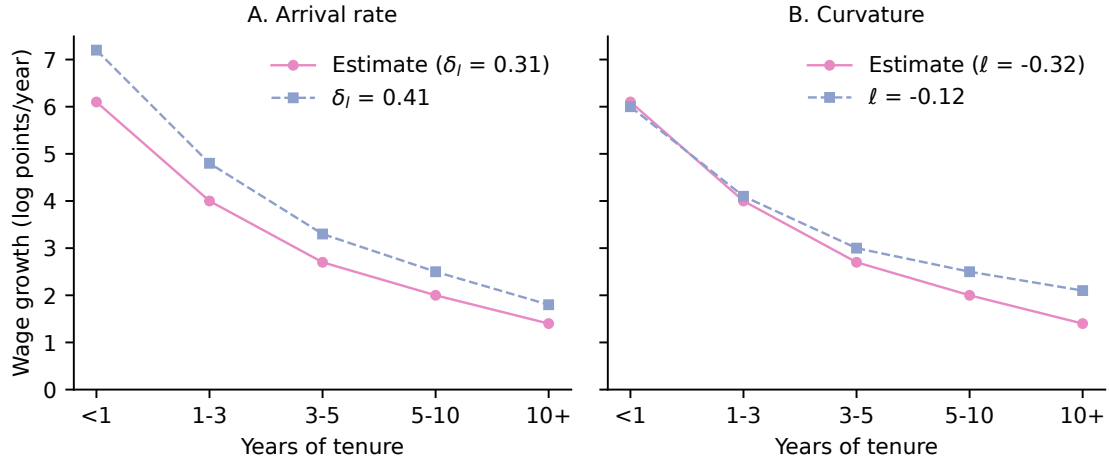


FIGURE 2. Identifying variation: the learning technology.

NOTES: Each statistic is computed on panels simulated from the model at the estimated parameters (solid line) and after moving one parameter, the others held at their estimates (dashed line); the legends give the value of the parameter moved. Mean annual change in log hourly wages within continuing matches, in log points, by years of tenure, in the bins of Figure 5.

Usage intensities from starting wages and displacement losses. The joint distribution of firms' usage intensities of portable and firm-specific human capital, ϕ_p and ϕ_s , is characterized by their means, μ_p and μ_s , their dispersions, σ_p and σ_s , and their correlation, ρ_ϕ . Identification exploits a key difference between the two types of capital: all employers value portable human capital from the start of a match, whereas firm-specific human capital is valuable only to the current employer and is lost when the match ends.

Portable human capital affects starting wages through two opposing forces. The *learning effect* is strongest for workers with little portable human capital. These workers have the most to gain from joining a firm that uses portable human capital intensively, because a larger share of the human capital they acquire will also be valued by future employers. They are therefore willing to accept lower starting wages in exchange for this learning opportunity. The *compensation effect* works in the opposite direction. As workers accumulate portable human capital, the expected gains from further learning decline, but their existing stock remains valuable to every employer. Firms that use portable human capital intensively must therefore offer higher starting wages to attract workers with more of it. Consequently, among inexperienced new hires, firms with higher ϕ_p pay lower starting wages, whereas among experienced new hires, they pay higher starting wages. As experience increases, the learning effect weakens and the compensation effect becomes stronger.⁷

The experience profile of starting wages identifies the distribution of ϕ_p . I divide new hires into five bins according to their experience at hire: less than one year, one to three years, three to five years, five to ten years, and ten or more years. I target the mean log starting wage in each bin relative to

⁷This is a local argument that requires human capital accumulation to diminish as the stock grows, as it does in the estimated model.

the first bin, as well as the within-bin variance of log starting wages. The mean profile disciplines μ_p : the greater the average intensity with which firms use portable human capital, the stronger both effects and the steeper the wage profile across experience bins. The within-bin variance disciplines σ_p : greater heterogeneity across firms in their use of portable human capital generates more variation in starting wages among otherwise comparable new hires. I measure this variance within experience bins because the two effects operate in opposite directions across experience levels. Pooling workers across bins would mix these opposing effects and obscure the wage variation attributable to heterogeneity in ϕ_p .

Firm-specific human capital is identified from the wage losses of workers displaced by firm closures. A worker displaced by a firm closure loses the firm-specific capital accumulated at the closing firm, whereas a newcomer at the same firm has accumulated little of it. In the panel, I classify a displaced worker as an incumbent if she was employed at the closing firm in the preceding year, and as a newcomer otherwise. For each worker and horizon, I measure the wage loss as the wage in her first job in the calendar year of that horizon minus her wage at the closing firm, net of the wage growth of stayers over the same years (Appendix A). The difference between incumbents' and newcomers' losses two years after the closure identifies μ_s .

The dispersion σ_s is identified using the difference between the variance of incumbents' losses and that of newcomers' losses one year after the closure, restricting the sample to workers re-employed in the same occupation. For incumbents, I measure losses relative to the mean loss in their tenure bin. Holding the other parameters fixed, a larger σ_s increases this variance gap. Incumbents differ in the amount of firm-specific capital accumulated in their jobs, whereas newcomers have little of it to lose. Greater heterogeneity across jobs in the amount of firm-specific capital they generate therefore translates into greater heterogeneity in incumbents' losses relative to newcomers' losses. Restricting the comparison to workers re-employed in the same occupation removes the dispersion in losses arising from the occupation haircut κ applied to portable capital. Subtracting the mean loss for each tenure bin also removes the increase in mean losses with tenure, which is governed by μ_s .

Finally, ρ_ϕ is identified from the relationship between displacement losses and the starting wage a worker earned upon joining the closing firm, and from how this relationship varies with experience at hire. Among workers who joined with little experience, a low starting wage signals a job with a high ϕ_p ; among workers who joined with substantial experience, a high starting wage does. If ϕ_p and ϕ_s are positively correlated, jobs with high ϕ_p also use firm-specific human capital intensively, so their workers have more to lose when the match ends. Displacement losses should then be larger following a low starting wage among inexperienced hires and following a high starting wage among experienced hires. If the correlation is negative, the pattern reverses.

I regress the wage loss one year after the closure on the starting wage, separately in each of the five bins of experience at hire, among displaced incumbents with one to seven years of tenure (starting

wages are observed only for matches that began in 2008 or later, which excludes longer tenures). The slope is negative in every bin, because a high starting wage also reflects a productive closing firm, and every displaced worker loses that. This component is common to all bins. The component due to ϕ_p is not: with a positive correlation the slope becomes more negative as experience at hire rises, with a negative correlation less negative. The change in the slope across experience bins therefore identifies ρ_ϕ .

Figure 3 summarizes the identification of the five parameters that govern the joint distribution of firms' usage intensities. Each panel plots a targeted statistic at the estimated parameters and after moving one parameter of the distribution of usage intensities, the others held at their estimates. Raising μ_p steepens the experience profile of mean starting wages (Panel A); raising σ_p raises the variance of starting wages within every bin, most among experienced hires (Panel B); and reversing the sign of ρ_ϕ reverses the change of the slope of the loss on the starting wage across experience at hire, which steepens instead of flattening (Panel C). Raising μ_s widens the gap between incumbents' and newcomers' mean losses (Panel D), and raising σ_s the gap between their variances (Panel E).

4.3. Estimated parameters and fit

Table 2 reports the estimates.

Search, bargaining and costs. Matching efficiency is identified by the rate at which unemployed workers find jobs ($A = 2.232$). This rate is 0.998 per year in the data and 1.070 in the model, so the average unemployed worker takes roughly a year to find a job. Conditional on A , the job-to-job transition rate identifies the relative search intensity of employed workers, who meet firms at η times the rate of unemployed workers. The estimate, $\eta = 0.283$, implies that employed workers search at slightly more than one-quarter of the intensity of the unemployed. The resulting job-to-job transition rate is 0.098 in the model, compared with 0.107 in the data, corresponding to approximately one in ten employed workers changing employers each year.

Bargaining power determines how the output of a match is divided between the worker and the firm, without affecting the size of the match surplus. It is therefore identified by the worker's share of output. In the data, this share is measured as the worker's annual wage (hourly wage multiplied by 1,820 hours) relative to value added per worker at her firm; in the model, it is the wage relative to match output, scaled by one minus the capital share of value added. The estimated bargaining power of workers is slightly above one-half ($\beta = 0.561$), yielding a mean output share of 0.525 in the model, compared with 0.502 in the data. The flow value of non-employment, $b = 0.712$, captures the combined value of unemployment benefits, home production, and leisure. It amounts to 58% of the mean wage. This parameter is identified from the lower tail of the distribution of wages in jobs entered from unemployment. A higher b raises workers' reservation wages, leading them to reject low-paying offers

TABLE 2. Estimated parameters and targeted moments.

Parameter	Symbol	Estimate	Identifying Moment	Data	Model
<i>Search and matching, prices and costs</i>					
Matching efficiency	A	2.232	U→E rate	0.998	1.070
Rel. search intensity	η	0.283	Job-to-job rate	0.107	0.098
Vacancy cost	c_v	0.040	Vacancies per firm	0.092	0.086
Entry cost	c_e	0.746	Firms per worker	0.983	0.906
Bargaining power	β	0.561	Worker's share, mean	0.502	0.525
Home production	b	0.712	Starting wages out of U, p50–p10	0.265	0.255
<i>Separation</i>					
Firm destruction	δ_e	0.158	E→U rate	0.129	0.146
Worker exit	δ_w	0.059	Mean experience (years)	14.1	13.6
<i>Occupational mobility</i>					
Cross-occ. meeting prob.	π_o	0.622	U→E rate, new occupation	0.537	0.494
Portable HC retention	κ	0.613	Displacement loss, new occupation	−0.134	−0.118
<i>Productivity process</i>					
Shock arrival	δ_z	0.434	ΔVA /worker within match, p90–p10	0.456	0.434
Shock dispersion	σ_z	0.718	ΔVA /worker within match, variance	0.055	0.049
<i>Production</i>					
Productivity weight	θ	0.566	VA/worker within match, autocorrelation	0.911	0.882
<i>Learning</i>					
Opportunity arrival	δ_l	0.308	Within-match wage growth	0.031	0.027
Increment curvature	ℓ	−0.321	Wage growth, tenure 10+	0.015	0.014
<i>Pre-market skill</i>					
Pre-market skill shape	a_α	0.935	Var. log wages, experience 0–1	0.062	0.062
Pre-market skill shape	a_β	0.300	Asym. log wages, experience 0–1	−0.126	−0.112
<i>Firm's usage intensity of human capital</i>					
log ϕ_p mean	μ_p	−1.469	Mean starting wage by experience at hire [†]	0.131	0.142
log ϕ_p s.d.	σ_p	0.974	Var. starting wage by experience at hire [†]	0.058	0.061
log ϕ_s mean	μ_s	−1.145	Displacement loss, incumbents – newcomers	−0.087	−0.079
log ϕ_s s.d.	σ_s	0.513	Var. displacement loss, incumbents – newcomers	0.035	0.033
Intensity correlation	ρ_ϕ	−0.301	Slope of loss on starting wage by experience at hire [†]	0.135	0.118

NOTES: Moments are computed identically in the panel data and in the simulated panel. The discount rate and the matching elasticity are fixed at $r = 0.05$ and $\mu = 0.5$ (Section 4.1). A and c_e are recovered from the equilibrium conditions (Appendix D). The worker's share is the hourly wage times 1,820 hours over value added per worker in the data, and the wage over match output scaled by 0.65, one minus the capital share of value added, in the model (Appendix E). The standard deviation of the wage measurement error is $\sigma_{me} = 0.081$, set by the autocorrelation of within-match wage changes: −0.232 in the data and −0.237 in the model. Tenure and experience are grouped into five bins: less than one year, one to three, three to five, five to ten, and ten or more years. The pre-market skill rows use the first experience bin and the ℓ row the last tenure bin. [†]One point of a targeted profile across the five bins of experience at hire: the ten-or-more bin relative to the first for the starting wage, the five-to-ten bin for its variance, and the ten-or-more bin minus the first for the slope; Figure 6 shows the full profiles. A starting wage is the wage of a match's first full-time year on a standard contract. Displacement losses are in log units and net of the wage growth of stayers; the incumbent–newcomer gap and the loss of workers re-employed in a new occupation are measured two years after the closure, the incumbent–newcomer variance gap (same-occupation re-employment, incumbents net of their tenure-bin mean) and the slope one year after.

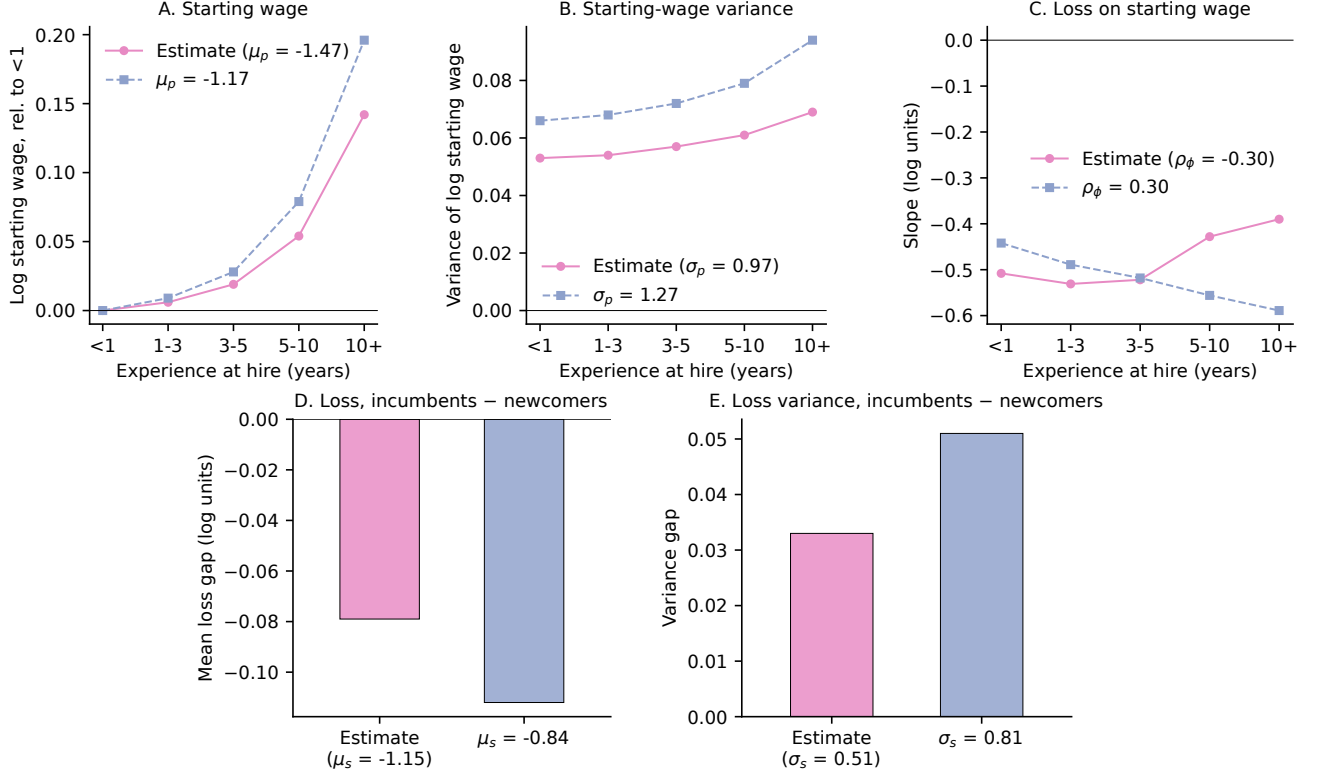


FIGURE 3. Identifying variation: the targeted statistics as one parameter of the distribution of usage intensities moves.

NOTES: Each statistic is computed on panels simulated from the model at the estimated parameters (solid line, first bar) and after moving one parameter of the distribution of usage intensities, the others held at their estimates (dashed line, second bar); the legends and labels give the value of the parameter moved. Panels A to C: the three experience-at-hire profiles of Figure 6. Panel D: the mean loss of displaced incumbents minus that of newcomers two years after the closure. Panel E: the variance of incumbents' losses minus that of newcomers' one year after the closure, among workers re-employed in the same occupation, incumbents net of their tenure-bin mean (Section 4.2).

and thereby thinning the lower tail. Accordingly, identification relies on the gap between the median and the 10th percentile of starting wages, which is 0.265 in the data and 0.255 in the model.

Firms enter until the expected value of entry just covers the entry cost, linking c_e to the number of firms per worker. Similarly, the cost of maintaining a vacancy, c_v , is linked to the number of vacancies per firm. The model matches vacancies per firm of 0.086, against 0.092 in the data, and firms per worker of 0.906 and 0.983 in the model and data, respectively. The estimated annual vacancy cost is $c_v = 0.040$, while the entry cost is $c_e = 0.746$, about nineteen times larger and approximately half (51%) of the annual output of an average match.

Only the mean job-to-job transition rate is targeted in estimation. Figure 4 provides an informative assessment of the model's ability to reproduce mobility patterns across tenure and experience. The model captures the decline in job-to-job transitions with tenure after the first year of a match. This pattern is consistent with the accumulation of firm-specific human capital, which workers forfeit

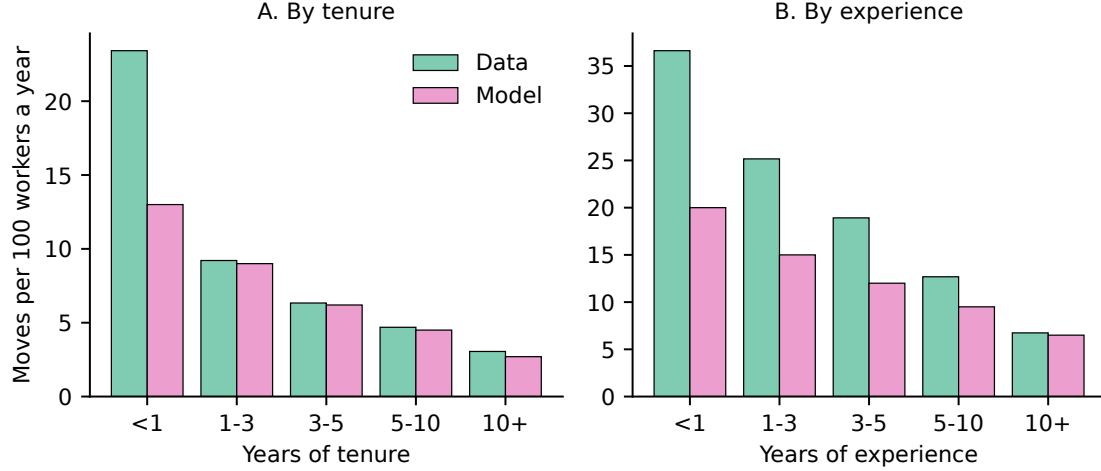


FIGURE 4. Job-to-job mobility by tenure and experience: data and model.

NOTES: Job-to-job moves per 100 employed workers a year, by years of tenure (Panel A) and of experience (Panel B). The model series are the same statistics computed on the panel simulated from the estimated model. None of the bins is targeted; the estimation targets the mean rate (Table 2).

when changing employers and which consequently reduces the attractiveness of outside offers. The model, however, substantially underpredicts mobility during the first year of a match, when workers have accumulated relatively little firm-specific human capital. A likely explanation is the prevalence of fixed-term contracts (Section 2), which can induce transitions to other employers when contracts expire.

The model also predicts declining job-to-job mobility with experience, but the decline is less pronounced than in the data. Two features of the model account for this discrepancy. First, productivity is redrawn approximately every two years, limiting how much a worker's position on the job ladder depends on how long she has been climbing it. Second, the model assumes that employed workers' search intensity is constant over the life cycle, whereas the data suggest that younger workers search more intensively. These features limit the model's ability to reproduce the steeper decline in mobility with experience observed in the data.

Separation. A firm is destroyed about once every six years ($\delta_e = 0.158$), and a worker leaves the labor market about once every seventeen years ($\delta_w = 0.059$). The destruction rate is set by the employment-to-unemployment rate once the productivity process has determined how many matches a productivity shock ends (Appendix E); the rate is 0.129 in the data and 0.146 in the model. The exit rate is set by mean experience, 14.1 years in the data and 13.6 in the model.

Occupational mobility. Most meetings cross occupations ($\pi_o = 0.622$): the share of hires out of unemployment into a new occupation is 0.537 in the data and 0.494 in the model. A cross-occupation move keeps about three fifths of the portable stock ($\kappa = 0.613$). For the average employed worker, giving up

two fifths of her portable stock would lower the output of her match by about 4.5 log points; the toll scales with the stock, so it deters precisely the experienced. The moment behind κ is the wage loss of displaced workers re-employed in a new occupation, -0.134 in the data and -0.118 in the model: a closure is an involuntary separation, so the separation is not selected on the value of switching, and the loss reflects the portable capital the move destroys on top of what every displaced worker loses.

Productivity and production. Firm productivity is redrawn about once every two years ($\delta_z = 0.434$), and each draw is large: with $\sigma_z = 0.718$ and a weight $\theta = 0.566$ on productivity in output, a draw one standard deviation away from the mean moves the log output of the match by 0.41. The three parameters are read off the dynamics of value added per worker within matches. The spread of its annual changes between the 10th and the 90th percentiles separates rare large shocks from frequent small ones, because a year without a shock sits at zero (0.456 in the data, 0.434 in the model); given the frequency, the variance of those changes gives the size (0.055 and 0.049); and the autocorrelation of value added per worker gives the weight, since the larger the part of output that each shock renews, the less persistent output is (0.911 and 0.882).

Learning. Learning opportunities arrive approximately once every three years ($\delta_l = 0.308$), and the associated productivity increments are substantial. For an entrant with average pre-market skill, the first learning opportunity raises match output by approximately 22 log points in the job held by the typical worker. The learning technology exhibits negative curvature ($\ell = -0.321$), so the expected increment declines with the accumulated composite. For the average employed worker, it is 14% smaller than for an entrant, implying that learning slows as workers accumulate human capital.

The arrival rate is identified by mean within-match wage growth, which is 0.031 per year in the data and 0.027 in the model. The curvature is identified by wage growth at ten or more years of tenure, 0.015 in the data and 0.014 in the model. Figure 5 compares the model and data profiles of within-match wage growth by tenure and experience. Only mean wage growth and the final tenure bin are targeted; the remaining bins therefore provide an out-of-target assessment of the learning technology's implications for wage dynamics.

The model reproduces the decline in wage growth with tenure, as diminishing increments reduce the pace of human capital accumulation as stocks grow. It also captures the hump-shaped profile of wage growth across experience, which reflects the combination of large early learning increments and wage gains triggered by outside offers as workers climb the job ladder. However, both profiles are flatter than in the data. Wage growth is substantially underpredicted during the first year of a match and early in a career. Since mean wage growth is targeted, this shortfall is offset by higher predicted growth at intermediate tenures.

As noted in Section 2, fixed-term contracts may help explain this discrepancy. When a worker's fixed-term contract is converted to an open-ended contract with the same employer, the transition may be

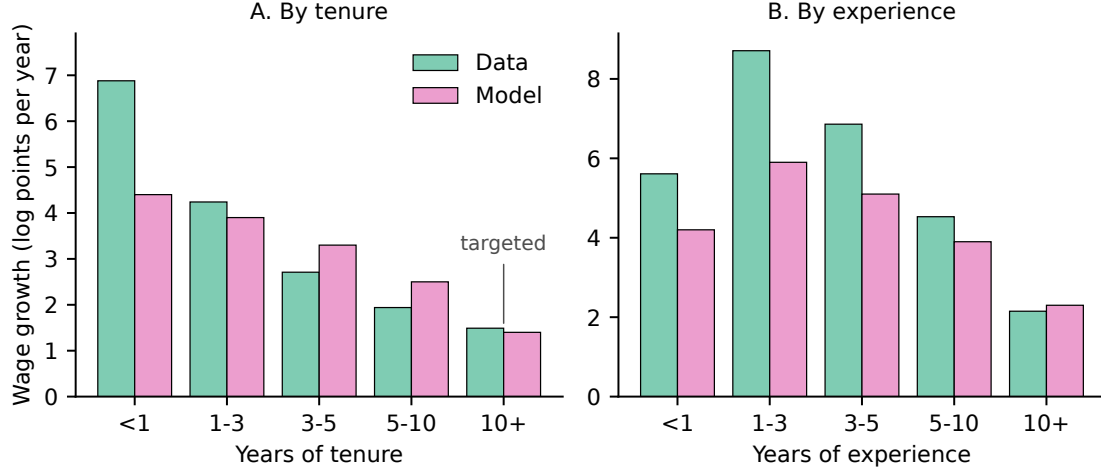


FIGURE 5. Within-match wage growth by tenure and experience: data and model.

NOTES: Mean annual change in residual log hourly wages within continuing matches, in log points, by years of tenure (Panel A) and of experience (Panel B), in the bins of Table 2. The model series are the same statistics computed on the panel simulated from the estimated model. Only the mean over all matches and the ten-or-more tenure bin, marked, are targeted in the estimation.

accompanied by an early-match wage increase. Because young workers hold a disproportionate share of fixed-term contracts, this mechanism could also contribute to the underprediction of wage growth early in the career.

Accumulated human capital constitutes a substantial share of the productive capacity of the average worker: it accounts for 36% of her composite $a + \phi_p p + \phi_s s$. This magnitude underscores the importance of distinguishing between the types of human capital accumulated across jobs, since the nature of that accumulation has direct implications for wage dynamics and the quantitative analysis in Section 6.

Pre-market skill. Both shape parameters of the Beta distribution of pre-market skill are below one ($a_\alpha = 0.935$, $a_\beta = 0.300$), and the second far below: the density rises toward the top of the support, so many entrants bring a high level of skill and a minority bring much less, and the spread of entry wages comes from that lower tail. The two parameters are read off log wages in the first year of experience, before learning and the job ladder have acted: their variance (0.062 in the data, 0.062 in the model) and their asymmetry, the gap between the 90th and the 50th percentiles minus the gap between the 50th and the 10th (−0.126 and −0.112), which is negative at entry because the lower tail is long. The asymmetry at later experience is not targeted, and its sign reverses in both the data and the model: as careers unfold, learning and the ladder stretch the upper tail, and among workers with ten or more years of experience the asymmetry is 0.063 in the data and 0.058 in the model.

Firms' usage intensities. The final block reports the parameters at the core of the paper. Three features of the estimates stand out. First, the median job allocates a larger share of each learning increment to firm-specific than to portable human capital ($\mu_s > \mu_p$). Second, heterogeneity across jobs is greater

in the intensity with which they use portable human capital than in their use of firm-specific human capital ($\sigma_p > \sigma_s$). Third, the two intensities are mildly negatively correlated ($\rho_\phi = -0.301$), indicating that jobs that use portable human capital more intensively tend to use firm-specific human capital less intensively.

Three of the five parameters are identified from profiles across experience at hire, rather than from individual moments (Section 4.2). Table 2 reports one point from each profile, while Figure 6 compares the full profiles in the data and the model. Panel A plots mean log starting wages by experience at hire, relative to workers hired with less than one year of experience. Panel B plots the variance of log starting wages within each experience bin. Panel C plots, by experience at hire, the slope of displaced incumbents' wage losses one year after a firm closure on their starting wages at the closing firm, restricting the sample to workers with one to seven years of tenure. Every bin in each profile is targeted in estimation.

The model reproduces all three profiles. Mean starting wages increase with experience at hire, with a particularly steep rise in the final bin, as governed by μ_p . The variance of log starting wages remains relatively flat across the initial bins before rising in the final bin, consistent with the dispersion in ϕ_p . Finally, the slope of displacement losses on starting wages is negative in every bin, since a higher starting wage also reflects the productivity of the closing firm. This slope becomes flatter among workers hired with more experience, providing evidence consistent with the estimated negative correlation between the two usage intensities.

The two displacement moments that identify μ_s and σ_s are also closely matched. The gap between incumbent and newcomer displacement losses is -0.087 in the data and -0.079 in the model; the corresponding gap in the variances of their losses is 0.035 in the data and 0.033 in the model. In both cases, the model matches the data to within a tenth of the data's value.

Figure 7 compares these targeted moments with additional displacement outcomes that are not targeted in estimation. Two years after a closure, the model reproduces the wage losses of incumbents and newcomers separately, rather than merely matching the difference between them. It also captures the losses of workers re-employed in the same occupation (Panel A). Panel B shows that losses rise with tenure at the closing firm, relative to non-displaced workers with the same tenure, occupation, experience and wage. The model reproduces this increase up to ten years of tenure: firm-specific human capital accumulates at the closing job, at a rate governed by μ_s , and is lost at the closure. Beyond ten years the model's losses fall short of the data's. Seniority premia set by collective agreements may account for part of the difference: they raise pay with tenure independently of skills, they are lost at a closure, and residualizing wages does not remove them, since it absorbs differences in levels across industries and occupations but not the return to tenure within them.



FIGURE 6. Targeted experience profiles: data and model.

NOTES: Bins of experience at hire, in years. Panel A: mean residual log starting wage relative to the first bin. Panel B: variance of residual log starting wages. Panel C: slope of the wage loss one year after the closure, net of the wage growth of stayers, of displaced incumbents with one to seven years of tenure on their residual log starting wage at the closing job, in log units. A starting wage is the wage of a match's first full-time year on a standard contract (Appendix A). The model series are the same statistics computed on the panel simulated from the estimated model.

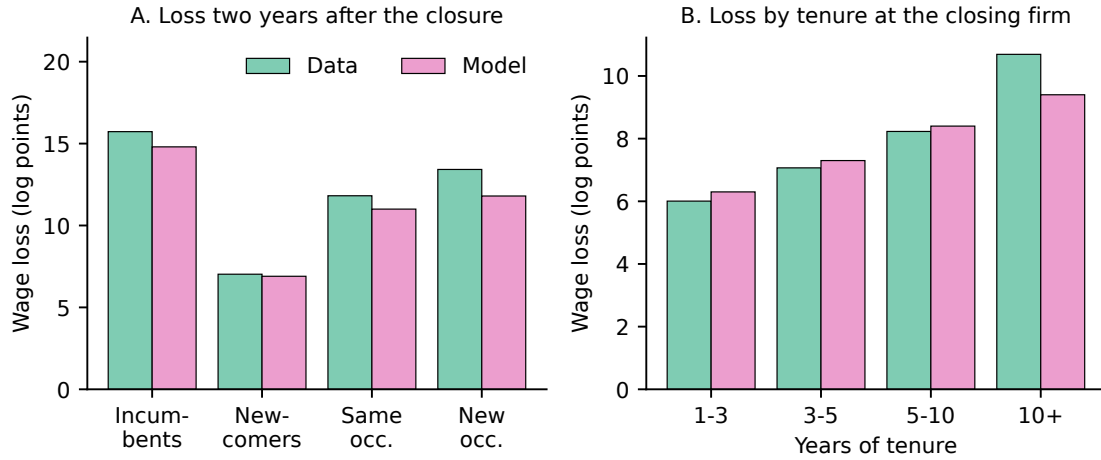


FIGURE 7. Displacement losses: data and model.

NOTES: Workers displaced by a firm closure. Panel A: mean wage loss two years after the closure, net of the wage growth of stayers, in log points, for incumbents, newcomers, and workers re-employed in the same or in a new occupation (Appendix A); only the incumbent–newcomer gap and the loss in a new occupation are targeted. Panel B: wage loss relative to matched controls averaged over two to four years after displacement, by years of tenure at the closing firm, from the event study of Section 6.2; not targeted. The model series are the same statistics computed on the panel simulated from the estimated model.

4.4. The learning mix across jobs

The first question this paper set out to answer is whether firms within an occupation differ in the portability of the human capital they build. Wages alone cannot answer this question. Wage growth with tenure and experience reflects several forces, including human capital accumulation and outside offers that bid up wages. Consequently, wage growth at a given job does not isolate the human capital

accumulated there. Moreover, workers are not randomly assigned to jobs: the matches that form and survive depend on workers' pre-market skills and on the portable and firm-specific human capital they have accumulated at previous jobs.

The estimated model recovers differences in the composition of learning across jobs through the five parameters in the final block of Table 2. At each learning opportunity, the share of the human capital increment that is portable is $\Delta(\phi_p, \phi_s)$, while the remaining share, $1 - \Delta(\phi_p, \phi_s) = \phi_s/(\phi_p + \phi_s)$, is firm-specific. Thus, the composition of learning at a job depends only on its two usage intensities, ϕ_p and ϕ_s , and not on the characteristics of the worker holding it.

Figure 8, Panel A, plots the joint density of log usage intensities across entering jobs implied by the estimates. The firm-specific share is constant along each ray from the origin because it depends only on the ratio of the two intensities. A job's distance from the origin along its ray, by contrast, indicates the intensity with which it uses human capital in production. The dotted rays correspond to firm-specific shares of one-quarter, one-half, and three-quarters; the dashed ray represents the mean share across filled jobs. The identification strategy in Section 4.2 determines the location of this density. Starting-wage profiles identify the horizontal dimension through μ_p and σ_p , while displacement losses identify the vertical dimension through μ_s and σ_s . Once both dimensions are identified, so is the firm-specific share of each job, given by the ray it lies on.

The estimated density is more dispersed along the portable-human-capital axis than along the firm-specific axis ($\sigma_p > \sigma_s$) and is tilted against the rays ($\rho_\phi < 0$). Together, these features imply substantial variation in the composition of learning across jobs. Panel B shows the implied distribution of firm-specific shares, and Table 3 reports summary statistics across filled jobs. At the median job, 58% of each learning increment is firm-specific, while the mean share is 55%. There is considerable heterogeneity around these central values: the firm-specific share is 23% at the 10th percentile and 86% at the 90th percentile.

Differences across jobs are more pronounced in the type of human capital they build than in the amount they generate. For the average employed worker, annual human capital accumulation at a job at the 90th percentile of the amount distribution is 1.2 times that at a job at the 10th percentile, whereas the firm-specific share at the 90th percentile (86%) is several times that at the 10th (23%).⁸ Because a job's composition of learning is the same regardless of who holds it, the estimated distribution characterizes heterogeneity across jobs rather than differences in the characteristics of their workers.

These results establish that jobs within the same occupation differ substantially in the type of human capital they build. Section 6 examines the wage implications of this heterogeneity using a counterfactual economy in which every job is moved to the dashed ray in Panel A, equalizing the firm-specific share

⁸The amount of human capital accumulated depends on a job's usage intensities interacting with the worker's existing stocks. At the start of a career, therefore, the amount does not vary across jobs. The composition of learning, by contrast, depends only on the job's usage intensities and is independent of the worker.

TABLE 3. The learning mix across jobs.

Firm-specific share of learning	
10th percentile	23%
Median	58%
90th percentile	86%
Mean	55%

NOTES: Distribution across jobs in the stationary equilibrium of the baseline economy, with jobs weighted by their mass of filled positions. The firm-specific share of a job is the share of each learning increment that it allocates to firm-specific human capital, $1 - \Delta(\phi_p, \phi_s)$.

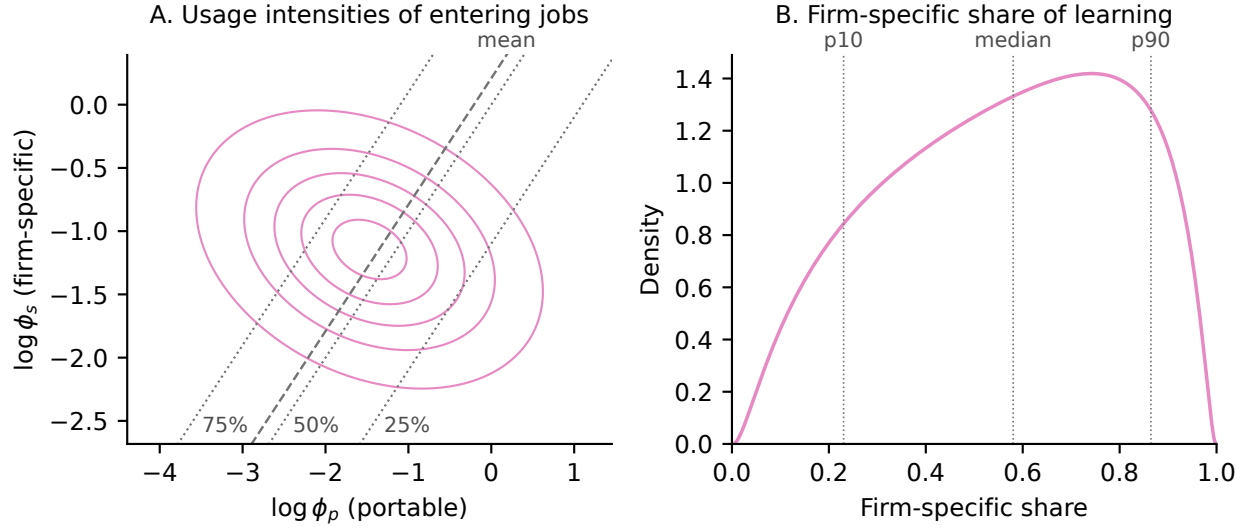


FIGURE 8. The estimated distribution of usage intensities and the learning mix across jobs.

NOTES: Panel A: contours of the joint density of $\log \phi_p$ and $\log \phi_s$ across entering jobs, the bivariate normal distribution with the estimated means, standard deviations and correlation of Table 2. Along each dotted ray the firm-specific share of learning, $\phi_s/(\phi_p + \phi_s)$, is constant at the value shown; the dashed ray is the mean share across filled jobs, 55%. Panel B: the density of the firm-specific share across entering jobs implied by the same distribution; the vertical lines mark its 10th percentile, median and 90th percentile across filled jobs in the stationary equilibrium (Table 3).

across jobs while preserving each job's distance from the origin.

5. Two Facts on Output and Wage Growth

This section evaluates the estimated model against two empirical facts that were not targeted in the estimation. Both bear on the paper's central mechanism: the first on whether part of what workers learn on the job is firm-specific, the second on whether jobs differ in how much of what they build is firm-specific.

The analysis proceeds in two steps. First, I assess how well the estimated model reproduces both facts. Second, I examine whether firm-specific human capital is essential to matching them. To this end, I compare the estimated model with a restricted version that excludes firm-specific human capital, making all learning portable across jobs. The restricted model retains the other mechanisms, including the job ladder and productivity shocks, and is re-estimated using the same targets (Appendix F). This comparison isolates the contribution of firm-specific human capital to the model's ability to account for the two empirical facts, holding the other mechanisms constant.

5.1. Fact 1: incumbents capture less of their contribution to output

The first fact compares a firm's newcomers, the workers who were not at the firm the year before, with its incumbents, those who were. It asks how a newcomer's output and wage compare with those of the average incumbent. While wages are observed, individual contributions to firm output are not. I therefore infer these contributions from changes in a firm's output per worker and average wage as its workforce composition shifts between incumbents and newcomers.

Each annual file of the employment register records every position at every firm in two consecutive years, allowing me to identify which workers are newcomers and which are incumbents in each firm-year. The balance sheets provide firm-level value added, while the employment register provides workers' earnings and hours. Using these data, I construct three firm-year variables. The first is the *newcomer share*, n_{ft} , defined as the share of firm f 's workforce in year t accounted for by newcomers, with each worker counted in proportion to her hours, that is, in full-time equivalents (FTE). The second is output per worker, y_{ft} , defined as the firm's value added divided by its total FTEs. The third is the mean log wage, $\bar{w}_{ft}$, defined as the FTE-weighted average of workers' log hourly wages.

Let y_{ft}^I denote the output per FTE among firm f 's incumbents in year t , and let y_{ft}^N denote the corresponding measure for newcomers. Neither is observed directly, but both are linked to observed output per worker through an accounting identity: total firm output is the sum of incumbents' and newcomers' output, so output per worker is their share-weighted average,

$$y_{ft} = (1 - n_{ft}) y_{ft}^I + n_{ft} y_{ft}^N. \quad (22)$$

Define the *incumbent output premium*, $P_{y,ft} = \log y_{ft}^I - \log y_{ft}^N$, as the log-point difference between the output of the average incumbent and that of the average newcomer. Appendix G shows that, to a first-order approximation in the premium, equation (22) can be expressed in logs as

$$\log y_{ft} \approx \log y_{ft}^I - P_{y,ft} n_{ft}. \quad (23)$$

Thus, log output per worker is approximately equal to the log output of the average incumbent minus a composition penalty proportional to the newcomer share and the incumbent output premium.

To illustrate, suppose each group's output remains constant while the firm replaces one-tenth of its incumbents with newcomers, increasing the newcomer share by ten percentage points. If each newcomer produces $P_{y,ft}$ log points less than an incumbent, this change reduces output per worker by approximately one-tenth of the premium. More generally, holding the premium fixed, a change in the newcomer share from year $t - 1$ to year t , $\Delta n_{ft} = n_{ft} - n_{f,t-1}$, changes log output per worker by approximately $-P_{y,ft} \Delta n_{ft}$.

Similarly, the mean log wage of the workforce is

$$\bar{w}_{ft} = (1 - n_{ft}) \bar{w}_{ft}^I + n_{ft} \bar{w}_{ft}^N = \bar{w}_{ft}^I - P_{w,ft} n_{ft}, \quad (24)$$

where $\bar{w}_{ft}^I$ and $\bar{w}_{ft}^N$ denote the mean log wages of the firm's incumbents and newcomers, respectively, and $P_{w,ft} = \bar{w}_{ft}^I - \bar{w}_{ft}^N$ is the firm's *incumbent wage premium*: the difference between the mean log wages of incumbents and newcomers.

Productive firms tend to produce more per worker, pay higher wages, and have lower turnover. Consequently, across firms, a low newcomer share is associated with high output per worker and high wages, even if incumbents and newcomers are equally productive and equally paid within each firm. To separate this cross-firm correlation from the relationship between workforce composition and firm outcomes, I difference both equations between consecutive years. Under the assumption that each firm's premium is constant between the two years, the change in log output per worker is

$$\Delta \log y_{ft} \approx \Delta \log y_{ft}^I - P_{y,f} \Delta n_{ft}, \quad (25)$$

where $P_{y,f}$ is the firm's premium. The wage equation in (24) can be differenced in the same way. Differencing removes persistent differences across firms, including differences in their underlying productivity, because these do not change from one year to the next. Subtracting each firm's mean over the whole sample would also remove them, but it would require the premium to be constant over up to seven years rather than two, and the newcomer share in every year to be unrelated to the productivity shocks of every other year, which fails if firms hire after a good year. Differencing consecutive years also lets the change in the mean wage be read from a single annual file of the register, which holds both years on the same definitions (Appendix A). It does not, however, remove year-to-year fluctuations in productivity. If a firm has a better year than the previous one, for example, its incumbents may produce more, and this increase remains in $\Delta \log y_{ft}^I$.

The change in incumbents' output, $\Delta \log y_{ft}^I$, is unobserved and therefore enters the regression of $\Delta \log y_{ft}$ on Δn_{ft} as part of the error term. Least squares identifies the premium if this error is uncorrelated with the change in the newcomer share. In other words, when a firm's newcomer share changes, its incumbents' output must not, for other reasons, change systematically with it. Differencing removes the main source of correlation between the two variables, persistent differences in productivity

across firms, but does not eliminate all potential sources of bias. Two remaining sources of correlation could cause the regression to attribute changes in incumbents' output to the newcomer share.

The first source of bias is year-to-year variation in firm productivity, which remains after differencing. Some of this variation is common to firms in the same industry. I absorb these common shocks using industry-by-year fixed effects at the two-digit industry level. Firm-specific productivity shocks, however, remain unobserved and constitute the main threat to the identifying assumption. If firms hire more when their prospects improve, an increase in the newcomer share coincides with an increase in incumbents' output. In that case, output per worker falls by less than the composition effect alone would predict, making the estimated coefficient on the change in the newcomer share less negative and the estimated output premium smaller. This bias therefore works against finding an output premium. It also works against finding a wedge: if wages respond less than output to a positive firm-specific shock, the wage premium is attenuated less than the output premium, biasing their difference downward as well.

The second source of bias is capital dilution. When a firm hires more workers, its existing capital is spread across a larger workforce, reducing output per worker, including the output of incumbents, independently of newcomers' productivity. I control for this effect using the change in the firm's log tangible fixed assets per FTE.

Finally, I include a control that serves a different purpose: rather than addressing a source of bias, it adjusts the comparison between incumbents and newcomers. Newcomers tend to be younger than incumbents, and experience raises both productivity and wages. The premiums would therefore partly reflect returns to experience rather than differences associated with being new to the firm. I control for the change in the workforce's mean age so that the estimated premiums compare incumbents and newcomers of similar ages.

Writing Δx_{ft} for the two controls, and $d_{y,qt}$ and $d_{w,qt}$ for the effects of industry $q = q(f)$ in year t , the equations I estimate are

$$\Delta \log y_{ft} = -P_y \Delta n_{ft} + \gamma'_y \Delta x_{ft} + d_{y,qt} + \varepsilon_{y,ft}, \quad (26)$$

$$\Delta \bar{w}_{ft} = -P_w \Delta n_{ft} + \gamma'_w \Delta x_{ft} + d_{w,qt} + \varepsilon_{w,ft}, \quad (27)$$

where $\varepsilon_{y,ft}$ denotes the change in incumbents' output not explained by the industry-year effects and controls, and $\varepsilon_{w,ft}$ is defined analogously for their wages. The error terms also absorb changes in each firm's premium between the two years; these changes are subject to the same identifying condition (Appendix G). The output and wage premiums are the negatives of the coefficients on the change in the newcomer share in the first and second equations, respectively.

The *output-wage wedge* is the difference between the two premiums, $\omega = P_y - P_w$: how much of the extra output of an incumbent is not paid to her. Its value can be read from the two regressions, but not

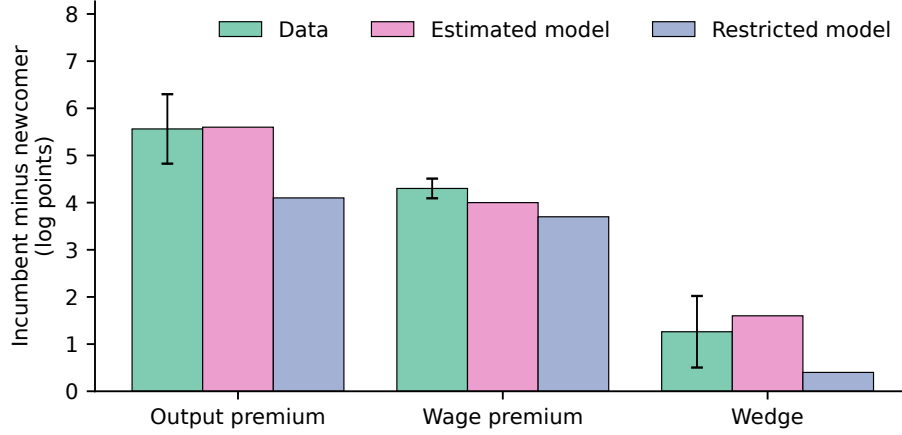


FIGURE 9. Fact 1: incumbents and newcomers in the data, the estimated model and the restricted model.

NOTES: Log points. Data: minus the coefficient on the change in the firm's newcomer share in a regression of the change in its log value added per full-time-equivalent worker (output premium), in its mean log hourly wage (wage premium), or in the difference between the two (wedge), with the controls of Section 5.1, on 784,751 firm-year pairs; whiskers give 95 percent confidence intervals, with standard errors clustered by firm. Estimated and restricted model: the gaps between incumbents and newcomers at the same firm state and with the same potential experience, computed directly in the simulated panels (Appendix G). Table A2 reports the estimates.

its standard error, because the two premiums are estimated on the same firms and their estimation errors are correlated. I therefore estimate the wedge directly. Let $g_{ft} = \log y_{ft} - \bar{w}_{ft}$ be the firm's output relative to its wages. Subtracting (27) from (26) gives

$$\Delta g_{ft} = -\omega \Delta n_{ft} + \gamma' \Delta x_{ft} + d_{qt} + \varepsilon_{ft}, \quad (28)$$

with $\gamma = \gamma_y - \gamma_w$, $d_{qt} = d_{y,qt} - d_{w,qt}$ and $\varepsilon_{ft} = \varepsilon_{y,ft} - \varepsilon_{w,ft}$. Because the regressors are identical, the least-squares coefficient in (28) is exactly minus the difference between the two estimated premiums, and its standard error is that of the wedge. In all three regressions, standard errors are clustered at the firm level, since the same firm contributes several pairs of years. Figure 9 reports the results, with the estimates and their standard errors in Table A2 of Appendix G.

An incumbent produces 5.6 log points more than a newcomer, but is paid only 4.3 log points more. The difference between the two gaps, the wedge, is 1.3 log points, with a standard error of 0.4. This is the first fact: an incumbent produces more than a newcomer, and her pay rises proportionally less than her output. I read the estimate as conservative: as argued above, the firm-specific shocks left in the error push it toward zero. If anything, the wedge in the data is larger.

In the model, the premiums do not need to be inferred. The regressions above are needed in the data because what each worker produces is not observed and firms differ in ways the econometrician does not see. In the model, both the output of each match and the state of each firm—its productivity and the intensities of its job—are observed. I therefore compute the premiums directly in the simulated

panel: the output premium is the gap in log output between incumbents and newcomers employed at firms in the same state and with the same potential experience, the years since entry into the labor market; the wage premium is the same gap in log wages; and the wedge is their difference. Comparing workers at firms in the same state is the model's counterpart of comparing them within the same firm, since two firms in the same state are identical. Appendix G gives the details and explains why the data's regressions, run on the model's one-worker firms, would not measure the same object.

The estimated model produces an output premium of 5.6 log points and a wage premium of 4.0, a wedge of 1.6 log points. It lies above the data's conservative estimate of 1.3 and within one standard error of it. The restricted model produces premiums of 4.1 and 3.7, a wedge of 0.4 log points, below the data's 95 percent confidence interval.

The wedge is the output a firm keeps, and a firm keeps only what no other firm would pay for. Portable skills are worth something at every firm, so outside offers price them: the firm keeps a share $1 - \beta$ of each increment only until an offer bids it away. Firm-specific skills are worth nothing elsewhere, so no offer prices them: the worker's share of them is the fraction β bargained when she learned them, and no later offer raises it. In the restricted model all skills are portable: an incumbent's extra output is bid into pay offer by offer, and the wedge is what the last offer has not yet reached.

5.2. Fact 2: returns to tenure and experience are weakly correlated across jobs

A job, as defined in Section 2, is the combination of a firm and an occupation. For each job, the object of interest is two numbers: the wage growth associated with one more year spent at that job—the return to tenure—and the wage growth associated with one more year spent in the labor market, at whatever employer—the return to experience. The natural starting point is a wage equation in which log wages depend on a match effect and match-specific slopes,

$$w_{i,t} = c_{i,j(i,t)} + \tau_{i,j(i,t)} \text{Ten}_{i,j(i,t),t} + \xi_{i,j(i,t)} \text{Exp}_{i,t} + \varepsilon_{i,t},$$

where $w_{i,t}$ is the residual log hourly wage of worker i in year t , $j(i,t)$ is the job employing her in that year, $\text{Ten}_{i,j(i,t),t}$ is her tenure—the number of years she has spent at that job—and $\text{Exp}_{i,t}$ is her experience—the number of years she has spent employed in the labor market. The coefficient $\tau_{i,j(i,t)}$ is the match's return to tenure, $\xi_{i,j(i,t)}$ is its return to experience, and $c_{i,j(i,t)}$ is a match fixed effect that absorbs everything about the pair that does not change over time.

This equation cannot be estimated, and the obstacle is identification rather than sample size. Within an employment relationship, tenure and experience advance one year for one year: a worker's experience is her experience at the start of the match, $\text{Exp}_{i,j}^0$, plus her tenure. Substituting $\text{Exp}_{i,t} = \text{Exp}_{i,j}^0 + \text{Ten}_{i,j,t}$

and writing j for $j(i, t)$, the equation reads

$$w_{i,t} = (c_{i,j} + \xi_{i,j} \text{Exp}_{i,j}^0) + (\tau_{i,j} + \xi_{i,j}) \text{Ten}_{i,j,t} + \varepsilon_{i,t}.$$

The first term in parentheses on the right-hand side does not change over the match, so the match fixed effect absorbs it, and what is left is a single regressor whose coefficient is the sum of the two returns: no quantity of data on one match separates $\tau_{i,j}$ from $\xi_{i,j}$. The separation has to come from across matches: from relationships that share a wage level and a pair of returns but began at different points of their workers' careers, since at the same tenure two such relationships differ only through ξ, Exp^0 . I therefore pool matches into classes: workers and jobs are each sorted into a small number of latent classes, estimated jointly with the coefficients, within which the level and the returns are common, in the spirit of the grouped fixed-effects estimator of Bonhomme and Manresa (2015). The restriction is that, within a class pair, experience at hire is unrelated to the part of match quality the class effects leave unexplained.

In practice, on all employed spell-years in my sample, I estimate

$$w_{i,t} = \zeta_{k(i)} + \psi_{g(j)} + \nu_{o(j)} + (\tau_{o(j)} + \tilde{\tau}_{k(i),g(j)}) \text{Ten}_{i,j,t} + (\xi_{o(j)} + \tilde{\xi}_{k(i),g(j)}) \text{Exp}_{i,t} + \varepsilon_{i,t}, \quad (29)$$

where $j = j(i, t)$ is the job employing worker i in year t , $k(i) \in \{1, \dots, K\}$ is the worker's latent class, $g(j) \in \{1, \dots, G\}$ the latent class of the job, and $o(j)$ the occupation of the job. Relative to the match-level equation, the estimator makes two changes. First, it coarsens matches into classes: the match fixed effect is replaced by a worker-class effect ζ_k and a job-class effect ψ_g , and the match-specific returns by a pair $(\tilde{\tau}_{k,g}, \tilde{\xi}_{k,g})$ shared by all matches between workers of class k and jobs of class g . Second, it gives every occupation its own wage level ν_o and its own returns, τ_o for tenure and ξ_o for experience. The reason is the question of the paper: the object of interest is how jobs differ *within* an occupation, and occupations may differ in how fast wages grow—wages may rise faster with tenure for engineers than for cashiers. With their own returns, those differences are absorbed by the occupation terms and cannot be mistaken for differences between jobs.

Two normalizations make the decomposition unique. Only relative wage levels are identified: adding the same constant to every worker-class effect and subtracting it from every occupation intercept leaves every predicted wage unchanged. So the first worker class and the first job class have an effect of zero, and the occupation intercepts carry the level of wages. The same is true of the returns: raising all cell deviations and lowering all occupation returns by the same amount changes nothing. So the deviations are set to average zero across cells, weighting each cell by its number of observations. The occupation returns then carry the level of the returns, and each deviation says how much faster or slower wages grow for a class of jobs than for the average job. For example, if the return to tenure is 0.8% per year among machinists and a class of jobs has a deviation of +0.5, a machinist in a job of that class gains 1.3% per year of tenure.

Equation (29) has two sets of unknowns: the coefficients, and the partition of workers and jobs into classes—which class each worker and each job belongs to. Each is straightforward to find given the other. Given the classes, the coefficients follow from a least-squares regression; given the coefficients, each worker and each job is best placed in the class that fits its wages best. For given (K, G) , the estimator alternates between the two, starting from an initial guess of the classes:

- a. *Initial classes.* Estimate a standard regression with a fixed effect for every worker and every job, and group workers with similar estimated effects into K classes and jobs with similar effects into G classes, by k -means.
- b. *Coefficients given the classes.* Estimate all coefficients in (29) by least squares, holding the classes fixed.
- c. *Classes given the coefficients.* Move each worker to the class that minimizes her own sum of squared residuals, holding the coefficients and all other assignments fixed; do the same for each job.
- d. Repeat steps b and c until the residual sum of squares stops improving.⁹

Because the algorithm jointly assigns workers and jobs to classes and estimates the returns associated with each class, using the same observations for both tasks risks overstating the dispersion in returns. For example, if a job's workers happen to experience unusually rapid wage growth in the sample, the algorithm may assign that job to a high-return class. Estimating the class's returns on those same observations would then partly reproduce the sampling noise that drove the assignment, making the estimated differences across classes appear larger than they truly are. To avoid this mechanical correlation, I separate classification from estimation. I randomly split the employed spell-years into five equally sized parts. In each rotation, I hold one part out, estimate the classes on the other four using the algorithm described above, then fix the resulting partition and re-estimate the coefficients of equation (29) by least squares on the held-out part. Spell-years involving a worker or job that does not appear in the classifying parts are excluded, since their class membership is unknown. Holding out each part in turn yields five partitions and five estimates of the class-specific returns, each based on observations not used to construct the corresponding partition. The statistics reported below average these five estimates, and the range across them is reported with each dispersion.

The worker classes are fixed at $K = 5$, both in the data and in the panel simulated from the model; they absorb differences in tenure and experience returns across workers, preventing these differences from being attributed to jobs. The number of job classes, G , sets the resolution at which jobs are told apart. With few classes, each pools jobs whose wages grow very differently, and averaging within the class hides the differences between them; as G grows, finer differences are resolved and the measured dispersion rises, until the classes are finer than the differences the data can separate and the dispersion levels off. I run the estimator at every class count in a grid from five to one hundred and read the

⁹Each step weakly reduces the residual sum of squares, so the algorithm converges, though not necessarily to a global minimum. The classification is therefore started from successive seeds of the k -means step until the lowest residual sum of squares has not improved over the last 10 starts, and the partition with the lowest residual sum of squares is kept.

dispersion against G (Figure 10). It levels off at $G^* = 40$ job classes, the count at which the statistics below are reported. Because the panel simulated from the model goes through the same procedure at every count, the two sides are measured at the same resolution and the comparison is like for like at each of them.

Estimation delivers one pair of deviations $(\tilde{\tau}_{k,g}, \tilde{\xi}_{k,g})$ for each cell, that is, for each combination of a worker class and a job class. To obtain the returns of a job class, I average its deviations over the worker classes employed in it, weighting each cell by its number of spell-year observations, $n_{k,g}$, and add back the average return across occupations, $\bar{\tau}$ for tenure and $\bar{\xi}$ for experience (the means of τ_o and ξ_o , weighted by spell-years):

$$\tau_g = \bar{\tau} + \sum_{k=1}^K \frac{n_{k,g}}{\sum_{k'} n_{k',g}} \tilde{\tau}_{k,g}, \quad \xi_g = \bar{\xi} + \sum_{k=1}^K \frac{n_{k,g}}{\sum_{k'} n_{k',g}} \tilde{\xi}_{k,g}. \quad (30)$$

Weighting each job class by its share of all spell-year observations gives the distribution of returns across job classes. Its location is the average return in the economy, while its dispersion is determined entirely by the within-occupation deviations: it captures how much the returns to tenure and experience differ across jobs within an occupation, not across occupations, which τ_o and ξ_o absorb.

The estimated dispersion depends on how accurately the procedure assigns jobs to classes. A job's class is inferred from its spell-years in the classifying parts of the sample. When a job is observed through only a few spell-years, its wage-growth pattern is estimated imprecisely and its assignment is uncertain. Such errors compress the estimated dispersion: jobs with different underlying returns land in the same class, and their differences are averaged away. The plateau of Figure 10 therefore measures the dispersion of returns at the resolution the observed job histories support; it need not equal the dispersion of the underlying job-specific returns, since some heterogeneity may remain unresolved. Appendix H examines how the estimate depends on the information used to classify jobs: forming the classes on one half, two thirds, four fifths or nine tenths of the rows yields the same dispersion once at least two thirds are used, so the plateau is not sensitive to this choice.

The comparison with the simulated panel also requires that jobs be observed with comparable information in the two samples. Under the same procedure, differences in the numbers of workers and spell-years observed per job change how accurately jobs are classified, and hence the measured dispersion. Appendix H therefore documents the structure of the job-level observations in the data and in the simulated panel, so that a difference in the two estimated dispersions can be read as a difference in wage-growth patterns rather than in how precisely jobs are classified. The panel simulated from the restricted model goes through the same procedure and is sampled by the same rule, so the three panels are measured alike.

Table 4 reports the mean and dispersion of the returns to tenure and experience across the $G^* = 40$ job classes, as well as the correlation between these returns across job classes. The statistics are reported

TABLE 4. The distribution of returns across jobs: data, estimated model and restricted model.

	Data	Full	Restricted
Mean return to tenure	0.58%	0.63%	0.60%
Mean return to experience	0.62%	0.57%	0.64%
p90–p10 of the return to tenure	2.00%	1.60%	1.30%
	[1.82%, 2.18%]	[1.46%, 1.74%]	[1.19%, 1.41%]
p90–p10 of the return to experience	0.80%	0.72%	0.95%
	[0.72%, 0.88%]	[0.66%, 0.78%]	[0.87%, 1.03%]
Correlation of the two returns	0.25	0.30	0.80
	[0.18, 0.31]	[0.24, 0.36]	[0.74, 0.85]

NOTES: Statistics of the distribution of job-class returns, in percent of the wage per year (log points), weighted by spell-years (class size), from the grouped fixed-effects estimator with $K = 5$ worker classes and $G^* = 40$ job classes. Cross-fitted: the spell-years are split at random into five parts; the classes are formed on four of them, started from successive seeds until the fit stops improving, and the returns are estimated on the fifth with the partition fixed; each part is held out once, the statistics average the five estimates, and brackets give their range across the rotations. The correlation is that of the job-class returns to tenure and to experience, weighted by class size. Full and Restricted: the same procedure on the panels simulated from the estimated model and from the restricted model, sampled as the data is.

for the data and for panels simulated from both the estimated and the restricted models. Figure 10 shows the two dispersions at every class count in the data and in the estimated model. In the data, the two returns are similar on average: a year of tenure raises the wage by 0.58% and a year of experience by 0.62%. They differ, however, in how much they vary across jobs. Within the same occupation, the return to tenure at the ninetieth percentile of job classes exceeds the tenth percentile by 2.00% per year, between 1.82% and 2.18% across the five rotations: over five years at the same job, that gap cumulates to about 10 percent of the wage. The return to experience varies across jobs by 0.80%. The figure shows how the two spreads are resolved as the classes become finer. The spread of tenure returns rises from 0.60% at five classes to 2.00% at G^* and is flat beyond it; the spread of experience returns is close to 0.80% at every count.

The estimated model's average returns are 0.63% for tenure and 0.57% for experience, and the restricted model's 0.60% and 0.64%. In both models, as in the data, the return to tenure varies across jobs more than the return to experience: the gaps between the ninetieth and tenth percentiles are 1.60% and 0.72% in the estimated model and 1.30% and 0.95% in the restricted model, against 2.00% and 0.80% in the data. The greater dispersion of tenure returns therefore does not separate the two models. In Figure 10, the estimated model's dispersions rise with G where the data's rise and are flat where the data's are flat, so the model resolves the differences between jobs at the same rate as the data does.

In the model, measured returns to tenure and experience reflect several forces at once: how much human capital jobs build and how portable it is, but also how firm productivity evolves, how outside offers bid wages up along the job ladder, and how bargaining splits the surplus of each match. None of

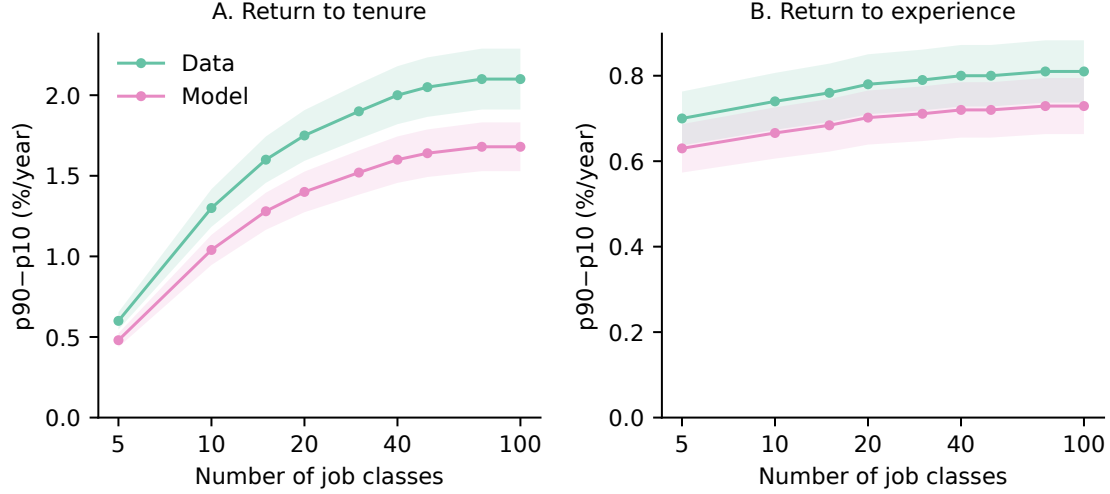


FIGURE 10. The dispersion of returns across job classes against the number of classes: data and model.

NOTES: Gap between the ninetieth and tenth percentiles of the job-class returns, in percent of the wage per year, weighted by spell-years, at every number of job classes from five to one hundred, with $K = 5$ worker classes; bands give the range over the five rotations of the cross-fitting. Panel A: the return to tenure. Panel B: the return to experience. Model: the same procedure on a panel simulated from the estimated model.

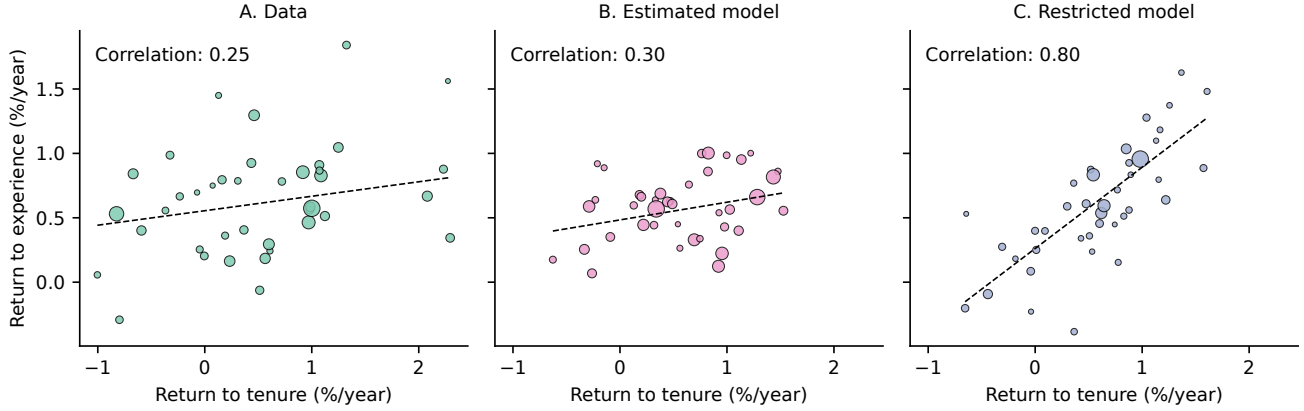


FIGURE 11. Fact 2: the returns to tenure and to experience across job classes.

NOTES: Each marker is one of the $G^* = 40$ job classes, at its return to tenure and its return to experience, in percent of the wage per year, averaged over the five rotations of the cross-fitting; marker area is proportional to the class's number of spell-years. The dashed line is the least-squares fit weighted by class size, and the correlation is the one of Table 4. Panels B and C: the same procedure on the panels simulated from the estimated and the restricted model, sampled as the data is.

the targeted moments describes how the returns vary across jobs, and the statistic that the two models move in opposite directions is the correlation of the two returns across job classes, the last row of Table 4. In the data it is 0.25, between 0.18 and 0.31 across the rotations: jobs that reward staying more reward time in the labor market only slightly more. The estimated model gives 0.30, between 0.24 and 0.36; the restricted model gives 0.80, between 0.74 and 0.85. Figure 11 shows the job classes behind the three correlations. In the restricted model, jobs differ in what they build only through ϕ_p , which raises both returns together. In the estimated model, ϕ_s raises the return to tenure without raising the

return to experience, which pulls the correlation down. The correlation is measured with noise, and because tenure and experience advance together within a match, the error in a class's two estimated returns biases every class-level correlation down; the three panels are sampled alike and go through the same procedure, so the comparison is between the panels rather than against zero. The two facts therefore separate the models in the same direction: without firm-specific human capital, the wedge falls below the data's confidence interval and the correlation of returns rises far above it.

6. Portability, Wage Inequality, and the Cost of Job Loss

This section answers the second question posed at the outset: how much the differences across jobs in the portability of the human capital they build contribute to wage inequality within occupations and to the wage cost of job loss. Section 4.4 measured those differences: the firm-specific share of learning averages 55% across jobs and ranges from 23% at the 10th percentile to 86% at the 90th (Table 3). To assess their consequences, I compare the baseline economy, the model at its estimated parameters, with a counterfactual economy in which all jobs use and build the two types of human capital in the same proportion. The comparison shows how much lower within-occupation wage dispersion and the wage cost of job loss would be without differences in the learning mix across jobs.

6.1. The learning mix and wage dispersion

To measure how much of within-occupation wage dispersion arises from differences across jobs in the composition of the human capital they use and build, I compare the baseline economy with a *common-mix* counterfactual economy. In the model economy, a job is characterized by the intensities with which it uses portable and firm-specific human capital, which in turn determine the amount of each type of human capital that workers accumulate on the job. In the common-mix counterfactual economy, the two usage intensities are proportional across jobs, so that all jobs use and workers accumulate the two types of human capital in the same proportion. Importantly, jobs in the common-mix counterfactual economy can still differ in the extent to which they build human capital; the counterfactual equalizes only the proportion in which each type of human capital is used and built across jobs.

The counterfactual economy is engineered as follows. In the model, a firm entering the economy draws its pair of usage intensities (ϕ_p, ϕ_s) from a joint log-normal distribution. The common-mix counterfactual economy changes that distribution and nothing else. Write the intensities with which a job uses portable and firm-specific human capital as $\phi_p = (1 - q)\Phi$ and $\phi_s = q\Phi$, where $\Phi = \phi_p + \phi_s$ measures how much overall human capital matters for production, and $q = \phi_s/(\phi_p + \phi_s)$ is the job's relative importance of firm-specific human capital in production. In the baseline economy both Φ and q vary across jobs. In the common-mix economy q is the same for every job, while the mean and the

dispersion of Φ are unchanged. The common share is the mean of the jobs' own shares in the baseline economy, 55% (Table 3).

Table 5 compares the dispersion of log wages in the baseline economy and in the common-mix economy. In the baseline economy the variance of log wages is 0.083. In terms of wages, a worker one standard deviation above the mean earns 33% more than a worker at the mean, and a worker two standard deviations above earns 78% more. In the common-mix economy the variance is 0.068, 18% lower than in the baseline economy. The gap between the 90th and the 10th percentiles decreases by only 6%: much of the fall in dispersion takes place in the tails of the wage distribution, the lower tail above all. Figure 12 plots selected percentiles of log wages in the two economies, relative to the median. Between the 10th and the 90th percentiles the two distributions differ little. Both tails are shorter in the common-mix economy, the lower tail by far the most: a worker at the 1st percentile earns 41% of the wage of a worker at the mean in the baseline economy, and 49% in the common-mix economy. The bottom 5 percent of wages account for 46% of the fall in the variance, and the top 5 percent for 22%.

Figure 13 shows who is affected. It sorts workers into five groups by pre-market skill, from the lowest to the highest, and reports the variance of log wages within each group in the two economies. Two patterns stand out. First, in both economies wage dispersion falls steeply with pre-market skill: in the baseline economy the variance is 0.342 in the lowest group and 0.048 in the highest. Second, the common-mix economy compresses wages far more among workers with low pre-market skill. The variance falls by 41% in the lowest group and by 21% and 16% in the next two (the three together hold 31% of entrants), and by less, 12% and 12%, in the two highest groups. Both patterns have the same source. Output depends on pre-market skill and on accumulated human capital through the composite $a + \phi_p p + \phi_s s$. For a worker with high pre-market skill the composite is dominated by a ; for a worker with low pre-market skill it is dominated by the human capital accumulated on the job. In the baseline economy, accumulated human capital contributes on average 77% of the composite in the lowest group and 31% in the highest. The wages of workers with low pre-market skill therefore depend mostly on what they accumulate at work and on how the job they hold uses it, and both differ across jobs; the wages of workers with high pre-market skill depend mostly on the skill they brought to the labor market, which does not. Differences across jobs in the mix of human capital they use and build are thus a much larger source of wage dispersion at the bottom of the pre-market skill distribution than at the top, and removing them compresses wages there the most.

6.2. The learning mix and the wage cost of job loss

The same experiment, applied to displaced workers, reveals how much of the wage cost of job loss stems from differences across jobs in the composition of the human capital they use and generate. I measure this cost using an event study around firm closures, applying the same procedure to the data and to panels simulated from the two economies. I define displaced workers as those who separate

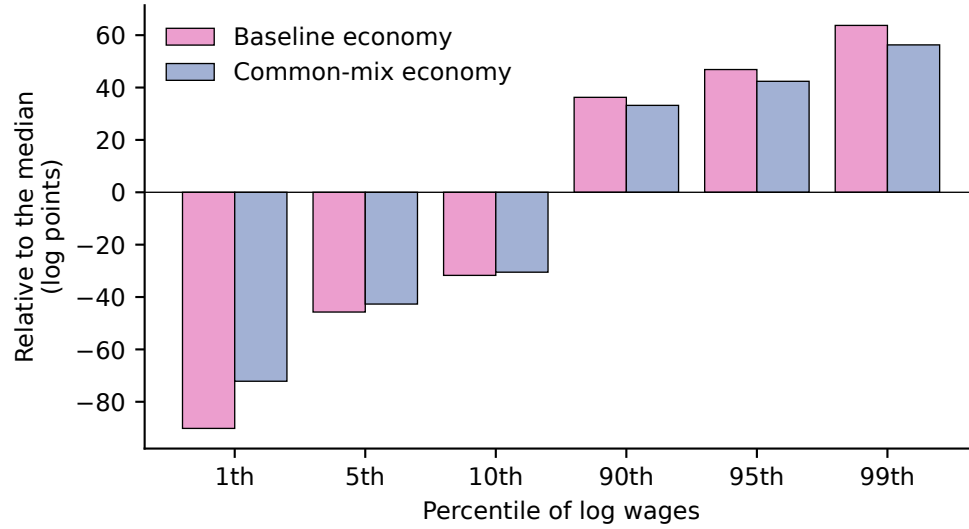


FIGURE 12. The distribution of log wages: baseline and common-mix economies.

NOTES: Percentiles of the log wage of employed workers in the two economies of Table 5, each minus the median of its own economy, in log points.

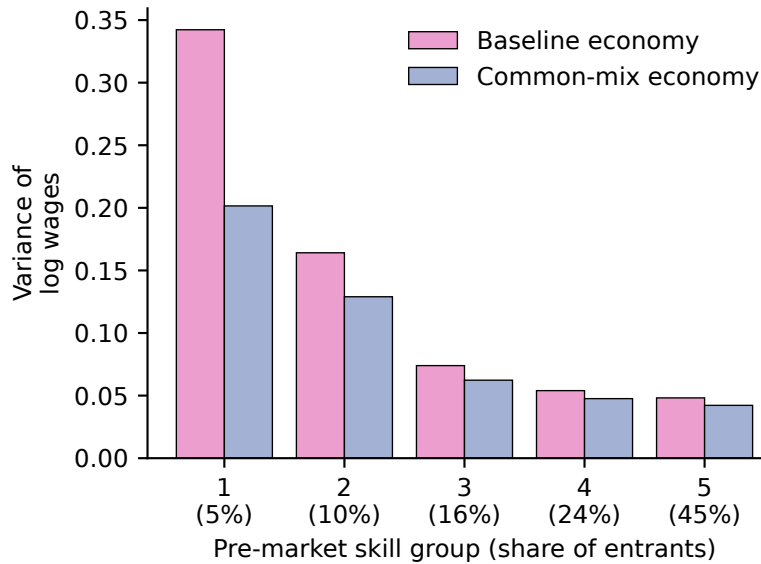


FIGURE 13. Wage dispersion by pre-market skill: baseline and common-mix economies.

NOTES: Variance of the log wage of employed workers in the two economies of Table 5, by group of pre-market skill a , from the lowest (1) to the highest (5). In parentheses: the group's share of labor-market entrants.

TABLE 5. Wage dispersion: baseline and common-mix economies.

	Baseline	Common mix	Change
Variance of log wages	0.083	0.068	–18%
p90–p10 of log wages	0.680	0.637	–6%

NOTES: Log hourly wages, computed on panels simulated from the two economies. In the common-mix economy every job has the same firm-specific share, 55%, the mean across jobs in the baseline economy; all other parameters are held at their estimated values (Section 6.1). Change is the common-mix value relative to the baseline value, minus one.

from a closing firm into unemployment either in the year of closure or in the preceding year. I refer to these two periods as the *event years* throughout this section. Each displaced worker with at least one year at the firm is matched to a worker who was never employed at a closing firm in one of its event years and who, in the year before the event, had the same occupation, tenure bin, experience bin, and within-occupation wage decile (Appendix A). The outcome is the gap in log hourly wages between the displaced worker after re-employment and her control. The loss I report is the average of this gap over the second, third and fourth years after displacement. I leave out the first year because it mixes the loss with the transition itself, when many of the displaced are still out of work or in a short first job; from the second year on, the gap captures the persistent part of the loss, and the fourth year is the last one observed for enough closures inside the sample window. The displaced are grouped by their tenure at the closing firm in the year of the event.

Figure 14 shows the loss by tenure in the baseline economy and in the common-mix counterfactual economy. In the baseline economy the loss rises with tenure, from 6.3 log points for workers with one to three years at the closing firm to 10.2 for workers with ten or more. In the common-mix economy the loss is smaller at every tenure, and the difference in log points grows with tenure: the loss is 20% smaller at one to three years, 5.0 log points, and 26% smaller at ten years or more, 7.5 log points. Over all displaced workers the loss is 23% smaller.

The wage loss is smaller in the counterfactual economy for the following reason. The contribution of firm-specific human capital to the output of a match, measured as in Panel B of Figure 15 by how much the log output of the match would fall without it, grows with the intensity with which the job uses firm-specific human capital and the stock the worker holds. A job with a higher firm-specific share raises both: it uses firm-specific human capital more intensively and, because it directs a larger share of every learning increment to it, its workers accumulate larger stocks. Intensity and stock therefore move together across jobs, and this co-movement raises the average contribution of firm-specific human capital to output. The common mix removes it, while leaving the average stock nearly unchanged. In the baseline economy, displaced workers in the third of jobs with the highest firm-specific share held 1.66 times the stock of the average displaced worker, and that stock contributed 2.20 times as much to output; in the third with the lowest share the figures are 0.47 and 0.17. Panel A of Figure 15 reports the average stock of firm-specific human capital that displaced workers held at the closing firm, and Panel



FIGURE 14. Displacement losses by tenure: baseline and common-mix economies.

NOTES: Wage loss relative to matched controls, in log points, averaged over two to four years after displacement, by tenure at the closing firm, measured on the simulated panels of the two economies by the event study of Section 6.2.

By the contribution of that stock to the output of the match; both are normalized to one for the average displaced worker of the baseline economy. At each tenure, displaced workers held nearly the same stock in the two economies, but the contribution of that stock to output was lower in the common-mix economy, by 22% over all displaced workers. Because wages reflect the output of the match, and a displaced worker loses her firm-specific human capital, a smaller contribution means a smaller wage loss.

In the estimated model, the rise of displacement losses with tenure at the closing firm is thus partly due to differences across jobs in the mix of human capital they use and build: removing them lowers the wage loss of displaced workers by 23% (20% at one to three years of tenure, 26% at ten or more), and flattens the rise of the loss with tenure by about 36%.

7. Conclusion

Firms are not just places where workers produce; they are also environments where they learn skills. But not all skills are the same. Since Becker's seminal work, labor economists have distinguished between general, or portable, human capital, which is valuable to any employer, and firm-specific human capital, which is valuable only at the firm where it was acquired. This paper demonstrates that to understand the role of firms as drivers of wage inequality and the wage cost of job loss, it is imperative to recognize that firms differ in how they use portable and firm-specific human capital,

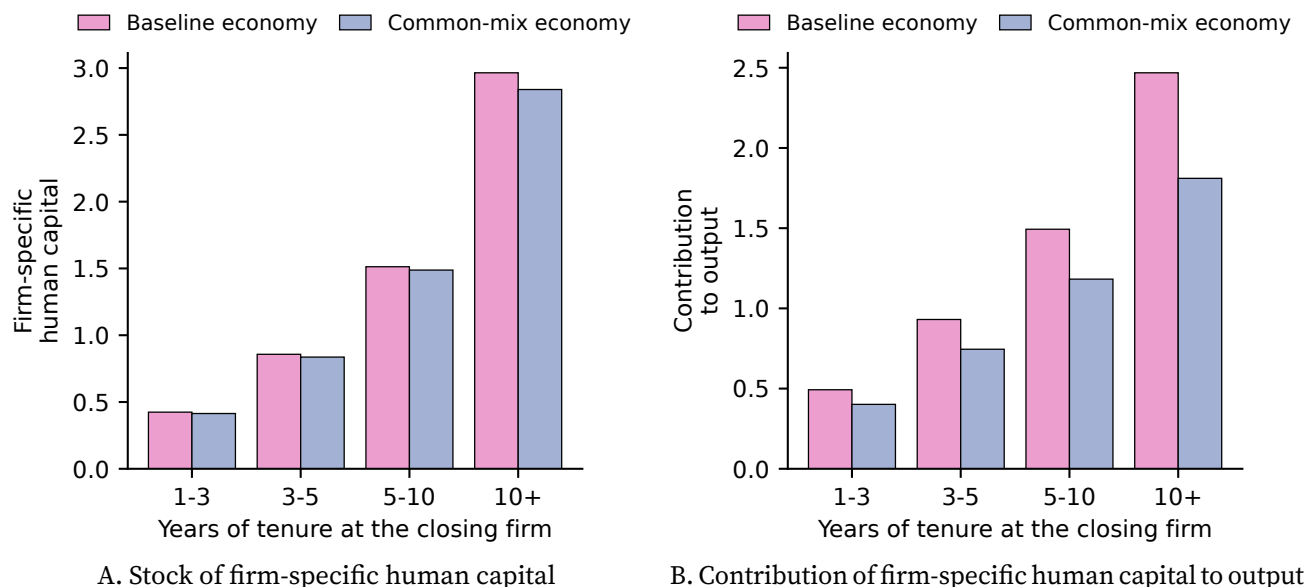


FIGURE 15. What displaced workers lose: baseline and common-mix economies.

NOTES: Simulated panels of Table 5; workers displaced by a closure after at least one year at the firm, by tenure at the closing firm. Panel A: the average stock of firm-specific human capital held at the closing firm in the last year there. Panel B: its contribution to output, the log output of the match minus the log output of the same match without that stock. Both are relative to their mean over all displaced workers of the baseline economy.

and how they equip workers with both types of skills.

I show the economic importance of differences in the portability of skills learned on the job by developing and estimating an equilibrium search model in which jobs differ in how intensively they use, and therefore build, portable and firm-specific human capital. The estimated model shows that jobs within the same occupation differ widely in the kind of human capital they build. On average, 55% of on-the-job skill learning is firm-specific, and across jobs the share runs from 23% at the tenth percentile to 86% at the ninetieth.

These differences matter for wages and for the wage cost of job loss. Without these differences in the share of firm-specific learning across jobs, the variance of within-occupation log wages would be 18% lower and the average wage loss from displacement would be 23% smaller. These effects are unevenly spread. Differences in the mix matter most for workers with the lowest pre-market skill, whose wages depend mostly on the human capital they accumulate at work and on how their job uses it, and for workers displaced after long tenure, who hold the most firm-specific human capital and lose disproportionately when their job relied heavily on it.

Several avenues of research point beyond the scope of this paper. In the model, firms are endowed with their usage intensities rather than choosing them; making them, and with them the specificity of training, a strategic choice would let the distribution itself respond to policy. Workers are risk-neutral, so the insurance value of a smaller displacement loss is not counted; risk aversion would give

the composition of learning a welfare dimension this paper does not claim. And occupations enter symmetrically, with occupation-specific human capital priced only through the loss of portable human capital on a move across occupations; letting employer-specific and occupation-specific human capital accumulate jointly would bring together two literatures that have so far been studied separately.

References

- Abowd, John M., Francis Kramarz, and David N. Margolis. 1999. "High Wage Workers and High Wage Firms." *Econometrica* 67 (2): 251–333. 10.1111/1468-0262.00020.
- Abowd, John M., Francis Kramarz, and Sébastien Roux. 2006. "Heterogeneity in Firms' Wages and Mobility Policies." In *Structural Models of Wage and Employment Dynamics*, vol. 275 of *Contributions to Economic Analysis*, edited by Henning Bunzel, Bent J. Christensen, George R. Neumann, and Jean-Marc Robin. Emerald Group Publishing. 10.1016/S0573-8555(05)75011-2.
- Adda, Jérôme and Christian Dustmann. 2023. "Sources of Wage Growth." *Journal of Political Economy* 131 (2): 456–503. 10.1086/721657.
- Adenbaum, Jacob, Fil Babalievsky, and William Jungerman. 2024. "Learning on the Job." Working paper, November 2024.
- Altonji, Joseph G. and Robert J. Shaker. 1987. "Do Wages Rise with Job Seniority?" *Review of Economic Studies* 54 (3): 437–459. 10.2307/2297568.
- Altonji, Joseph G. and Nicolas Williams. 2005. "Do Wages Rise with Job Seniority? A Reassessment." *ILR Review* 58 (3): 370–397. 10.1177/001979390505800304.
- Arellano-Bover, Jaime and Fernando Saltiel. 2026. "Differences in On-the-Job Learning across Firms." *Journal of Labor Economics* 44 (1): 149–188.
- Audoly, Richard, Federica De Pace, and Giulio Fella. 2022. "Job ladder, human capital, and the cost of job loss." *FRB of New York Staff Report* (1043).
- Bagger, Jesper, François Fontaine, Fabien Postel-Vinay, and Jean-Marc Robin. 2014. "Tenure, experience, human capital, and wages: A tractable equilibrium search model of wage dynamics." *American Economic Review* 104 (6): 1551–1596.
- Balke, Neele and Thibaut Lamadon. 2022. "Productivity Shocks, Long-Term Contracts, and Earnings Dynamics." *American Economic Review* 112 (7): 2139–2177. 10.1257/aer.20200302.
- Beffy, Magali, Moshe Buchinsky, Denis Fougère, Thierry Kamionka, and Francis Kramarz. 2006. "The Returns to Seniority in France (and Why Are They Lower than in the United States?)." IZA Discussion Paper 1935.
- Bonhomme, Stéphane and Elena Manresa. 2015. "Grouped Patterns of Heterogeneity in Panel Data." *Econometrica* 83 (3): 1147–1184. 10.3982/ECTA11319.
- Borovičková, Katarína and Claudia Macaluso. 2025. "Heterogeneous job ladders." *Journal of Monetary Economics* 150: 103711. 10.1016/j.jmoneco.2024.103711.
- Bowlus, Audra J. and Huju Liu. 2013. "The contributions of search and human capital to earnings growth over the life cycle." *European Economic Review* 64: 305–331. 10.1016/j.euroecorev.2013.10.002.
- Buchinsky, Moshe, Denis Fougere, Francis Kramarz, and Rusty Tchernis. 2010. "Interfirm mobility, wages and the returns to seniority and experience in the United States." *Review of Economic Studies* 77 (3): 972–1001.
- Burdett, Kenneth, Carlos Carrillo-Tudela, and Melvyn Coles. 2020. "The Cost of Job Loss." *Review of Economic Studies* 87 (4): 1757–1798. 10.1093/restud/rdaa014.
- Cahuc, Pierre, Fabien Postel-Vinay, and Jean-Marc Robin. 2006. "Wage Bargaining with On-the-Job Search: Theory and Evidence." *Econometrica* 74 (2): 323–364. 10.1111/j.1468-0262.2006.00665.x.
- Devereux, Paul J., Robert A. Hart, and J. Elizabeth Roberts. 2013. "Job Spells, Employer Spells, and Wage Returns to Tenure." IZA Discussion Paper 7384.
- Dustmann, Christian and Costas Meghir. 2005. "Wages, Experience and Seniority." *Review of Economic Studies* 72 (1): 77–108. 10.1111/0034-6527.00325.

- Engbom, Niklas. 2022. "Labor Market Fluidity and Human Capital Accumulation." NBER Working Paper 29698.
- Fallick, Bruce, John Haltiwanger, Erika McEntarfer, and Matthew Staiger. 2025. "Job Displacement and Earnings Losses: The Role of Joblessness." *American Economic Journal: Macroeconomics* 17 (2): 177–205. 10.1257/mac.20210329.
- Gregory, Victoria. 2021. "Firms as Learning Environments: Implications for Earnings Dynamics and Job Search." Working paper, Federal Reserve Bank of St. Louis.
- Herkenhoff, Kyle, Jeremy Lise, Guido Menzio, and Gordon M. Phillips. 2024. "Production and Learning in Teams." *Econometrica* 92 (2): 467–504. 10.3982/ECTA16748.
- Hernández-Martínez, Victor, Hans A. Holter, and Roberto Pinheiro. 2025. "The Hedgehog's Curse: Knowledge Specialization and Displacement Loss." *Labour Economics* 97.
- Hethcote-Maier, Tanja and Johannes F. Schmieder. 2013. "Does the Use of Worker Flows Improve the Analysis of Establishment Turnover? Evidence from German Administrative Data." *Schmollers Jahrbuch* 133 (4): 477–510. 10.3790/schm.133.4.477.
- Huckfeldt, Christopher. 2022. "Understanding the Scarring Effect of Recessions." *American Economic Review* 112 (4): 1273–1310. 10.1257/aer.20160449.
- Jacobson, Louis S., Robert J. LaLonde, and Daniel G. Sullivan. 1993. "Earnings Losses of Displaced Workers." *American Economic Review* 83 (4): 685–709.
- Jarosch, Gregor. 2023. "Searching for Job Security and the Consequences of Job Loss." *Econometrica* 91 (3): 903–942. 10.3982/ECTA14008.
- Jarosch, Gregor, Ezra Oberfield, and Esteban Rossi-Hansberg. 2021. "Learning from Coworkers." *Econometrica* 89 (2): 647–676. 10.3982/ECTA16915.
- Jung, Philip and Moritz Kuhn. 2019. "Earnings Losses and Labor Mobility over the Life Cycle." *Journal of the European Economic Association* 17 (3): 678–724. 10.1093/jea/jvy014.
- Karahan, Fatih, Serdar Ozkan, and Jae Song. 2019. "Anatomy of Lifetime Earnings Inequality: Heterogeneity in Job Ladder Risk vs. Human Capital." Federal Reserve Bank of New York Staff Report 908.
- Krolkowski, Pawel. 2017. "Job Ladders and Earnings of Displaced Workers." *American Economic Journal: Macroeconomics* 9 (2): 1–31. 10.1257/mac.20140064.
- Lachowska, Marta, Alexandre Mas, and Stephen A. Woodbury. 2020. "Sources of Displaced Workers' Long-Term Earnings Losses." *American Economic Review* 110 (10): 3231–3266. 10.1257/aer.20180652.
- Lise, Jeremy and Fabien Postel-Vinay. 2020. "Multidimensional Skills, Sorting, and Human Capital Accumulation." *American Economic Review* 110 (8): 2328–2376. 10.1257/aer.20162002.
- MacLeod, W. Bentley and James M. Malcomson. 1993. "Investments, Holdup, and the Form of Market Contracts." *American Economic Review* 83 (4): 811–837.
- Margolis, David N. 1996. "Cohort Effects and Returns to Seniority in France." *Annales d'Économie et de Statistique* (41/42): 443–464.
- Petrongolo, Barbara and Christopher A. Pissarides. 2001. "Looking into the black box: A survey of the matching function." *Journal of Economic Literature* 39 (2): 390–431. 10.1257/jel.39.2.390.
- Pires, Tiago, Arek Szydlowski, and Shuai Zhao. 2024. "Firm-specific Human Capital Accumulation: Evidence from Brazil." *Oxford Bulletin of Economics and Statistics* 86 (3): 568–605. 10.1111/obes.12570.
- Postel-Vinay, Fabien and Jean-Marc Robin. 2002. "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity." *Econometrica* 70 (6): 2295–2350. 10.1111/j.1468-0262.2002.00441.x.
- Rubinstein, Yona and Yoram Weiss. 2006. "Post-Schooling Wage Growth: Investment, Search and Learning." *Handbook of the Economics of Education* 1: 1–67. 10.1016/S1574-0692(06)01001-4.

- Schmieder, Johannes F., Till von Wachter, and Jörg Heining. 2023. "The Costs of Job Displacement over the Business Cycle and Its Sources: Evidence from Germany." *American Economic Review* 113 (5): 1208–1254. 10.1257/aer.20200252.
- Theodoropoulos, Nikolaos and Carl Singleton. 2026. "Returns to Tenure: A Critical Assessment of the Evidence and Interpretation." IZA Discussion Paper 18705.
- Topel, Robert H. 1991. "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority." *Journal of Political Economy* 99 (1): 145–176.
- Yamaguchi, Shintaro. 2010. "Job search, bargaining, and wage dynamics." *Journal of Labor Economics* 28 (3): 595–631.

Appendix A. Data Construction

This appendix records the construction steps summarized in Section 2.

Worker panel. The sample frame is workers born in October of even-numbered years, 1 in 25 of the salaried population; the odd-year October births and the birth-date supplement added to the raw panel from 2002 are dropped, so age parity equals calendar-year parity throughout. The window is 2008 to 2016: compensated unemployment spells are recorded from 2008, and from 2017 the employment records come from the monthly DSN declaration, whose reported hours are not on the same basis as the annual declaration's, so hourly wages are not comparable across the break. I retain private-sector employment, dropping public-sector employment categories, State and local-government employers, and the overseas departments; the sample covers mainland France. Workers are aged 16 to 65. I drop home workers, spells with a missing contract type, and spells with fictitious worker or firm identifiers. Employment spells are required to report between 455 and 4,160 paid hours, which excludes marginal attachments and implausible records. Cumulated unemployment within a spell sequence is capped at two years. Worker-years containing overlapping spells are dropped, as are worker-years with more than three sequential spells, since in both cases the spell sequence cannot be ordered unambiguously. Log hourly wages are trimmed at the first and ninety-ninth percentiles. Wage statistics use full-time employment spells on standard contracts. Earnings are deflated by the INSEE consumer price index with 2015 as the base year.

Employment relationships and occupation. An employment relationship is a maximal run of consecutive employment spells at the same firm; it ends with any change of firm or any spell out of employment. Occupation is assigned as the modal occupation observed across all spell-year observations within the relationship; in the event of a tie, the occupation with the smallest code is selected. If a relationship reports, say, code 543 in four of its five years and code 542 in one, all five of its years are assigned code 543. Among relationships observed in more than one year, 66.4% report a single code in every year, and the assignment alters the recorded code in 11.3% of employed spell-years. The code changes it overwrites have three sources. The first is a change of source: the panel takes the occupation code of a spell-year from one variable up to 2009 and from another from 2010, the second taking precedence whenever it is recorded, so a relationship observed on both sides of that year can change code although the worker's position does not change. The second is miscoding: isolated years that revert to the surrounding code the following year. The third is genuine change: a worker who is promoted or moves to a different position inside her firm can legitimately be recorded under a new code. The single-occupation rule deliberately does not follow these changes: for the analysis, a worker remains in the same job for as long as she remains with her firm.

Tenure. Tenure with the employer and labor-market experience are taken directly from the panel, which computes them from each worker's full employment history; tenure is measured at the level of

the firm. Because the single-occupation rule does not follow within-firm moves, tenure is the number of years a worker has spent with her employer, not the number of years she has spent in her current position; a promotion does not reset the clock. This choice avoids a selection problem. Firms promote their stronger performers out of positions, so the workers observed with long tenure in a single position are disproportionately those who were not promoted—weaker performers on average. Position tenure therefore carries information about who the worker is, not only about how long she has worked: Devereux, Hart, and Roberts (2013) show that among workers with the same employer tenure, those with longer position tenure earn *less*. Employer tenure is free of this selection, because promotions move workers between positions without ending the employment relationship.

Firm accounts. The balance-sheet files change variable naming conventions in 2009, when the current business-statistics system was introduced, and the table layout changes across vintages; the extraction maps each vintage’s variable codes onto a single set of names. Firm identifiers are standardized to nine digits and records that cannot be parsed are dropped. In the rare case of duplicate records for a firm-year, the record with the largest non-missing value added is retained. Value added is deflated by the same price index used for wages; employment is not deflated, and the worker’s share, a ratio of contemporaneous quantities, is unaffected.

Residualization. The projections of Section 2.5 are estimated on employed spells and applied to those same spells. Unemployment spells retain their recorded replacement income: industry, region and occupation are all undefined outside an employment relationship, and no statistic in the paper reads the unemployment wage. Region is taken from the worker’s establishment, industry is the two-digit division of the French activity classification, recorded at the firm, and occupation—in the wage projection only—is the three-digit PCS code used throughout the paper. The projections weight each spell-year by its duration, and each residual is re-centered so that its duration-weighted mean equals that of the original variable: the means of the log hourly wage and of log value added per worker are unchanged, and the transformation affects only the variation around them.

Closure classification. A firm is a candidate for death in year N if it has at least one non-marginal employment relationship with positive pay in year $N - 1$ and none in year N . For each candidate, every worker employed there in $N - 1$ is followed to her highest-paying employer in year N ; workers not employed in N are retained in the denominator. Let σ be the share of the candidate’s last-year workforce absorbed by the single largest successor employer. Firms with fewer than three employees in their last active year are left unclassified. Among the rest, $\sigma \leq 0.30$ is classified as a closure, $0.30 < \sigma \leq 0.50$ as ambiguous, and $\sigma > 0.50$ as a change of identifier or a takeover. Only closures are used. The implementation departs from Hethey-Maier and Schmieder (2013) in two respects: it operates on firms rather than establishments, and it examines only where workers go, not where they come from, so it does not separately identify spinoffs. Each closure has two event years: its last active year and the year before it, the run-up, in which the firm is already shedding workers.

Incumbents and newcomers. A newcomer at a firm in year N is a worker who held no position of at least 455 paid hours at that firm in $N - 1$; any other worker at the firm is an incumbent. Fact 1 applies this rule within one annual file of the register, which holds both years. The panel applies it to every displacement statistic, in the data and in the simulated panels, where it reads: no employment spell at the same firm in $N - 1$; the incumbent–newcomer gap that identifies μ_s is computed under this rule in the data and in the model. Displacements in the first year of the panel, 2008, cannot be classified because the preceding year is not observed; they enter the statistics on all displaced workers but neither group.

Fact 1. One annual file of the register holds every position in years $N - 1$ and N . A position counts if it is full time, on a standard contract (open-ended or fixed-term), in the private sector, with at least 455 paid hours and an hourly wage between 3 and 300 euros, and held by a worker aged 16 to 65. Positions are grouped by worker and firm, so a worker who moves between two establishments of the same firm stays at the firm. A newcomer in year N is a worker in the firm’s year- N workforce who held no position of at least 455 paid hours at the firm in $N - 1$, and the newcomer share is the newcomers’ share of the full-time equivalents of that workforce. The change in the newcomer share compares the shares read from two consecutive files. The change in the wage is the mean log hourly wage of the year- N workforce minus that of the year- $N - 1$ workforce, both weighted by full-time equivalents and read from the same file. The change in mean age is the same difference for age, with each worker’s age taken from the same file on both sides, so that it reflects only who left and who joined. Capital per worker is the balance sheet’s tangible fixed assets over its full-time equivalents. The files 2010 to 2016 are used, so the changes run from 2010–2011 to 2015–2016. The sample keeps the firm-year pairs in which the firm reports positive value added, at least ten full-time equivalents and positive tangible fixed assets in both $N - 1$ and N . The regressions of Table A2 are unweighted, include a dummy for each two-digit industry and year, and cluster standard errors by firm.

Displacement in the panel. A displacement is the last spell-year of an employment relationship at a closing firm in one of its event years, full time on a standard contract, directly followed by an unemployment spell. Every such episode counts, so a worker displaced by two closures contributes two. No tenure floor applies to the displaced population. A floor of one year of tenure applies only where a statistic is about the firm-specific capital incumbents have accumulated: the variance, the p90–p10 gap and the share of negative values of incumbents’ losses, the loss-on-starting-wage slopes, and the event study. The mean losses of incumbents and newcomers, the gap between them, and their re-employment shares use every episode. In the simulated panels, a firm closes when it dies through either channel of Section 3.3.4—exogenous destruction, or a productivity shock that drives the value of a vacancy negative. As in the data, the displaced are the workers whose relationship with the firm ends in the year it dies and who move into unemployment, whatever ended the relationship, and the same rules apply. Two differences between the two sides remain. First, the model has no run-up year: a firm’s death is a single event, so the simulated displaced all separate in the year the firm dies, whereas the

data also count those who separate in the year before. Second, the panel records unemployment only while it is compensated, whereas the simulated panel records every unemployment spell. A displaced worker in the data is therefore one who draws unemployment benefits after the separation; one who finds a job before claiming, or who is not entitled to benefits, is not counted. This selection has no counterpart in the model and affects every displacement statistic computed in the panel: the event study and the moments that identify μ_s , σ_s , κ and ρ_ϕ .

Displacement losses. The loss at horizon $h \in \{1, 2\}$ is the log hourly wage of the worker's first full-time spell-year on a standard contract in calendar year $L + h$, minus her log hourly wage at the closing firm in L , net of the wage growth of stayers from L to $L + h$: the sum over those years of the mean within-match wage change. Losses are reported as wage changes, so a larger loss is a more negative number. She counts as re-employed at horizon h if she holds any employment spell in $L + h$; her loss is measured when one of these spells is full time on a standard contract. Re-employment in the same or a different occupation compares the occupation of that spell-year with that of the closing job. A horizon past the last year of the panel is not measured: losses two years after are not measured for displacements in 2015, and the simulated panels apply the same rule at their own last year.

Starting wages. A match's starting wage is the log hourly wage of its first full-time spell-year on a standard contract, kept only when that spell-year falls within the first year of tenure, so that the start of the match is observed; experience at hire is the experience recorded in that spell-year. Each match counts once. A hire out of unemployment is a match whose first spell-year directly follows an unemployment spell of the same worker. Starting wages are therefore observed only for matches that start in 2008 or later. For this reason the loss-on-starting-wage slopes keep displaced incumbents with at most seven years of tenure at the closing firm, in the data and in the simulated panels. In each bin of experience at hire, the slope is that of a least-squares regression of the loss one year after the closure on the starting wage at the closing job.

Event study. The event study keeps, for each worker, her first displacement with at least one year of tenure. She is matched, with replacement, to one control, observed in the calendar year of her event, who shares her occupation, her one-year tenure bin (up to ten years, then ten to fifteen and fifteen or more), her experience bin (less than five years, five to fifteen, fifteen or more) and her within-occupation decile of the residual wage, all measured in the year before the event; within the cell, the control is the worker with the nearest residual wage. Controls exclude every worker ever employed at a closing firm in one of its event years; in the simulated panels, every worker employed at a dying firm in the year it dies or the year before. The gap in residual log hourly wages is followed from three years before to five years after the event, and the tenure groups of Section 6.2 use tenure at the event. Standard errors come from 500 bootstrap draws over closures: each draw resamples closing firms together with their displaced workers and keeps each worker's matched control.

Other measurement details. Within-match wage changes are differences in log hourly wages between consecutive spell-years of the same relationship, both full time on standard contracts. Within-match changes in log value added per worker are trimmed: pairs of consecutive years whose change lies below the first or above the ninety-ninth percentile are dropped before any statistic is computed on them, the autocorrelation of levels included. The grouped fixed-effects estimator of Section 5 is applied to a simulated panel of the size of the data: the first simulated workers, each taken whole, whose careers together contain as many employed spell-years with a wage as the data.

Appendix B. Value Functions

This appendix derives the gross flow value equations for employed workers and filled firms, and shows how the corresponding net flow values follow by subtracting the unemployment value (15) and the vacancy value (16), respectively: the worker's net value (17) is reported in the body, and the filled firm's net value (A3) is reported here. Throughout, $\Psi(\cdot)$ denotes the label-averaged effective surplus operator from (14), $\Delta = \Delta(\phi_p, \phi_s)$, and the regions $\mathcal{P}_1(\chi), \mathcal{P}_2(\chi), \mathcal{Z}_0, \dots, \mathcal{Z}_3, \mathcal{H}$ together with the bargained wages $w_1^\chi, w_2^\chi, w^+, w^-, w_3$ are as defined in Section 3.

B.1. Employed workers

The flow value of employment for a type- (a, p, s) worker matched with a type- (z, ϕ_p, ϕ_s) firm under a wage contract w is

$$\begin{aligned}
& rW_e(a, p, s, z, \phi_p, \phi_s, w) \\
&= \underbrace{w - \delta_w W_e(a, p, s, z, \phi_p, \phi_s, w)}_{\text{worker death}} + \underbrace{\delta_e [W_u(a, p) - W_e(a, p, s, z, \phi_p, \phi_s, w)]}_{\text{exogenous separation}} \\
&+ \eta \alpha \Xi_W(a, p, s, z, \phi_p, \phi_s, w) \\
&+ \delta_z \left[\int_{\mathcal{Z}_0} [W_u(a, p) - W_e(a, p, s, z, \phi_p, \phi_s, w)] \gamma_z(z') dz' \right. \\
&\quad + \int_{\mathcal{Z}_1} [W_e(a, p, s, z', \phi_p, \phi_s, w^+) - W_e(a, p, s, z, \phi_p, \phi_s, w)] \gamma_z(z') dz' \\
&\quad + \int_{\mathcal{Z}_2} [W_e(a, p, s, z', \phi_p, \phi_s, w^-) - W_e(a, p, s, z, \phi_p, \phi_s, w)] \gamma_z(z') dz' \\
&\quad \left. + \int_{\mathcal{Z}_3} [W_e(a, p, s, z', \phi_p, \phi_s, w) - W_e(a, p, s, z, \phi_p, \phi_s, w)] \gamma_z(z') dz' \right] \\
&+ \delta_l \int_{\mathcal{H}} [W_e(a, p + \Delta h, s + (1 - \Delta)h, z, \phi_p, \phi_s, w_3)
\end{aligned}$$

$$- W_e(a, p, s, z, \phi_p, \phi_s, w)] \gamma_h(h|a, p, s, \phi_p, \phi_s) dh, \quad (A1)$$

where the poaching term Ξ_W is given by (18). Subtracting the unemployed-worker HJB (15) from (A1) yields the net flow value of employment, equation (17) in the body. The term $-\alpha\beta \iiint \Psi(\cdot; 0) v$ in (17) is the unemployed worker's search value, carried over from (15) by the subtraction; the matched worker's own search gains enter through Ξ_W .

B.2. Filled firms

The flow value of production for a type- (z, ϕ_p, ϕ_s) firm matched with a type- (a, p, s) worker under wage contract w is

$$\begin{aligned} r\Pi_f(a, p, s, z, \phi_p, \phi_s, w) &= \underbrace{y(a, p, s, z, \phi_p, \phi_s) - w}_{\text{flow profit}} - \underbrace{\delta_e \Pi_f(a, p, s, z, \phi_p, \phi_s, w)}_{\text{firm death}} \\ &+ \underbrace{\delta_w [\max\{\Pi_v(z, \phi_p, \phi_s), 0\} - \Pi_f(a, p, s, z, \phi_p, \phi_s, w)]}_{\text{exogenous separation}} \\ &+ \eta\alpha \Xi_F(a, p, s, z, \phi_p, \phi_s, w) \\ &+ \delta_z \left[\int_{Z_0} [\max\{\Pi_v(z', \phi_p, \phi_s), 0\} - \Pi_f(a, p, s, z, \phi_p, \phi_s, w)] \gamma_z(z') dz' \right. \\ &+ \int_{Z_1} [\Pi_f(a, p, s, z', \phi_p, \phi_s, w^+) - \Pi_f(a, p, s, z, \phi_p, \phi_s, w)] \gamma_z(z') dz' \\ &+ \int_{Z_2} [\Pi_f(a, p, s, z', \phi_p, \phi_s, w^-) - \Pi_f(a, p, s, z, \phi_p, \phi_s, w)] \gamma_z(z') dz' \\ &\left. + \int_{Z_3} [\Pi_f(a, p, s, z', \phi_p, \phi_s, w) - \Pi_f(a, p, s, z, \phi_p, \phi_s, w)] \gamma_z(z') dz' \right] \\ &+ \delta_l \int_{\mathcal{H}} [\Pi_f(a, p + \Delta h, s + (1 - \Delta)h, z, \phi_p, \phi_s, w_3) \\ &- \Pi_f(a, p, s, z, \phi_p, \phi_s, w)] \gamma_h(h|a, p, s, \phi_p, \phi_s) dh, \quad (A2) \end{aligned}$$

where the firm-side poaching term Ξ_F is given by (A4) below. Subtracting the vacancy HJB (16) from (A2) yields the net flow value of production,

$$\begin{aligned} &(r + \delta_w + \delta_e) [\Pi_f(a, p, s, z, \phi_p, \phi_s, w) - \Pi_v(z, \phi_p, \phi_s)] \\ &= y(a, p, s, z, \phi_p, \phi_s) - w + c_v - \delta_z \int (\max\{\Pi_v(z', \phi_p, \phi_s), 0\} - \Pi_v(z, \phi_p, \phi_s)) \gamma_z(z') dz' \\ &- \alpha(1 - \beta) \iint \Psi(a', p', z, \phi_p, \phi_s; 0) u(a', p') da' dp' \end{aligned}$$

$$\begin{aligned}
& -\eta\alpha(1-\beta) \int \cdots \int \Psi(a', p', z, \phi_p, \phi_s; S(a', p', s', z', \phi'_p, \phi'_s)) \\
& \quad \times \Gamma(a', p', s', z', \phi'_p, \phi'_s) da' dp' ds' dz' d\phi'_p d\phi'_s \\
& + \eta\alpha \Xi_F(a, p, s, z, \phi_p, \phi_s, w) \\
& + \delta_z \left[\int_{\mathcal{Z}_0} [\max\{\Pi_v(z', \phi_p, \phi_s), 0\} - \Pi_f(a, p, s, z, \phi_p, \phi_s, w)] \gamma_z(z') dz' \right. \\
& \quad + \int_{\mathcal{Z}_1} [\Pi_f(a, p, s, z', \phi_p, \phi_s, w^+) - \Pi_f(a, p, s, z, \phi_p, \phi_s, w)] \gamma_z(z') dz' \\
& \quad + \int_{\mathcal{Z}_2} [\Pi_f(a, p, s, z', \phi_p, \phi_s, w^-) - \Pi_f(a, p, s, z, \phi_p, \phi_s, w)] \gamma_z(z') dz' \\
& \quad \left. + \int_{\mathcal{Z}_3} [\Pi_f(a, p, s, z', \phi_p, \phi_s, w) - \Pi_f(a, p, s, z, \phi_p, \phi_s, w)] \gamma_z(z') dz' \right] \\
& + \delta_l \int_{\mathcal{H}} [\Pi_f(a, p + \Delta h, s + (1 - \Delta)h, z, \phi_p, \phi_s, w_3) \\
& \quad - \Pi_f(a, p, s, z, \phi_p, \phi_s, w)] \gamma_h(h|a, p, s, \phi_p, \phi_s) dh, \tag{A3}
\end{aligned}$$

where the poaching term for the filled firm is

$$\begin{aligned}
\Xi_F(a, p, s, z, \phi_p, \phi_s, w) & = (1 - \pi_o) \iiint_{\mathcal{P}_1(0)} [\Pi_f(a, p, s, z, \phi_p, \phi_s, w_1^0) \\
& \quad - \Pi_f(a, p, s, z, \phi_p, \phi_s, w)] v dz' d\phi'_p d\phi'_s \\
& + (1 - \pi_o) \iiint_{\mathcal{P}_2(0)} [\max\{\Pi_v(z, \phi_p, \phi_s), 0\} \\
& \quad - \Pi_f(a, p, s, z, \phi_p, \phi_s, w)] v dz' d\phi'_p d\phi'_s \\
& + \pi_o \iiint_{\mathcal{P}_1(1)} [\Pi_f(a, p, s, z, \phi_p, \phi_s, w_1^1) \\
& \quad - \Pi_f(a, p, s, z, \phi_p, \phi_s, w)] v dz' d\phi'_p d\phi'_s \\
& + \pi_o \iiint_{\mathcal{P}_2(1)} [\max\{\Pi_v(z, \phi_p, \phi_s), 0\} \\
& \quad - \Pi_f(a, p, s, z, \phi_p, \phi_s, w)] v dz' d\phi'_p d\phi'_s. \tag{A4}
\end{aligned}$$

The Ψ -forms of the forgone-search terms in (A3) arise analogously to the worker case.

B.3. Match surplus

The surplus $S(a, p, s, z, \phi_p, \phi_s) = W_e(\cdot) - W_u(a, p) + \Pi_f(\cdot) - \max\{\Pi_v(z, \phi_p, \phi_s), 0\}$ is independent of the wage. Adding the net flow value equations (17) and (A3) eliminates w on the right-hand side and

produces, after the simplifications described below, the surplus fixed-point equation (19) in the body.

The simplifications proceed as in the canonical sequential-auction model, adapted to the cross-occupation labeling. Throughout this list, $S \equiv S(a, p, s, z, \phi_p, \phi_s)$ denotes the surplus of the current match. Specifically:

- a. In each region $\mathcal{P}_1(\chi)$, the wage renegotiation is a pure transfer between worker and firm, so the integrand $[W_e(w_1^\chi) - W_e(w)] + [\Pi_f(w_1^\chi) - \Pi_f(w)] = 0$ and these terms drop out.
- b. In each region $\mathcal{P}_2(\chi)$, the integrand simplifies to $\beta[\tilde{S}(\cdot; \chi) - S]$ by adding and subtracting $W_u(a, p)$. Since $\mathcal{P}_2(\chi) = \{\tilde{S}(\cdot; \chi) > S\}$, the integral over the full support with integrand $\beta \max\{\tilde{S}(\cdot; \chi) - S, 0\}$ recovers the same value. Averaging over χ with weights $1 - \pi_o$ and π_o produces $\eta\alpha\beta \iiint \Psi(a, p, z', \phi'_p, \phi'_s; S) v(z', \phi'_p, \phi'_s) dz' d\phi'_p d\phi'_s$.
- c. In region $\mathcal{Z}_0$, adding and subtracting $\max\{\Pi_v(z, \phi_p, \phi_s), 0\}$ rewrites the integrand as $\max\{\Pi_v(z'), 0\} - \max\{\Pi_v(z), 0\} - S$ (suppressing ϕ_p, ϕ_s).
- d. In regions $\mathcal{Z}_1, \mathcal{Z}_2, \mathcal{Z}_3$, adding and subtracting $W_u(a, p)$ and $\max\{\Pi_v(z, \phi_p, \phi_s), 0\}$ rewrites each integrand as $S(a, p, s, z', \phi_p, \phi_s) - S + \max\{\Pi_v(z'), 0\} - \max\{\Pi_v(z), 0\}$, since S is defined using $\max\{\Pi_v, 0\}$ in (5). Combining all four $\mathcal{Z}$ regions, the $\max\{\Pi_v, \cdot\}$ residuals integrate to $\int [\max\{\Pi_v(z'), 0\} - \max\{\Pi_v(z), 0\}] \gamma_z(z') dz'$, which cancels exactly against the $-\delta_z \int [\max\{\Pi_v(z'), 0\} - \Pi_v(z)] \gamma_z(z') dz'$ term in (A3); the equivalence uses $\max\{\Pi_v(z), 0\} = \Pi_v(z)$ at the current state, which holds on the support of Γ by the equilibrium definition. The result is the compact form

$$\delta_z \int [\mathbf{1}\{\Pi_v(z', \phi_p, \phi_s) \geq 0\} \max\{S(a, p, s, z', \phi_p, \phi_s), 0\} - S] \gamma_z(z') dz'.$$

The indicator is not redundant: on the part of $\mathcal{Z}_0$ where destruction is triggered by the firm's exit, the post-shock surplus may be positive, yet the pair separates and the continuation contributes zero.

- e. In the learning region $\mathcal{H}$, adding and subtracting $W_u(a, p + \Delta h)$, $W_u(a, p)$, and $\Pi_v(z)$ yields the integrand

$$S(a, p + \Delta h, s + (1 - \Delta)h, z, \phi_p, \phi_s) - S + W_u(a, p + \Delta h) - W_u(a, p).$$

The unmatched-search terms enter the surplus equation with a negative sign from both sides: the worker's forgone search value, $-\alpha\beta \iiint \Psi(a, p, z', \phi'_p, \phi'_s; 0) v(z', \phi'_p, \phi'_s) dz' d\phi'_p d\phi'_s$, and the vacancy's forgone value of hiring an unemployed worker, $-\alpha(1 - \beta) \iint \Psi(a', p', z, \phi_p, \phi_s; 0) u(a', p') da' dp'$.

Substituting and combining yields the surplus fixed-point equation (19) in the body.

Appendix C. Stationary Densities

This appendix states the stationarity conditions of Section 3: the balance equations that the densities Γ , u , and v and the entry rate μ_J satisfy in a stationary equilibrium.

To compactly handle the cross-occupation labeling in flow equations, define the label-conditional match-formation and switching indicators. For an unmatched worker at (a, p) meeting a vacancy at (z, ϕ_p, ϕ_s) ,

$$\begin{aligned}\Lambda(a, p, z, \phi_p, \phi_s) &= (1 - \pi_o) \mathbf{1}\{S(a, p, 0, z, \phi_p, \phi_s) \geq 0\} \\ &\quad + \pi_o \mathbf{1}\{S(a, \kappa p, 0, z, \phi_p, \phi_s) - L_u(a, p) \geq 0\}\end{aligned}$$

is the unconditional match-formation probability. For an employed worker at (a, p, s) in a firm at (z, ϕ_p, ϕ_s) meeting a vacancy at (z', ϕ'_p, ϕ'_s) ,

$$\begin{aligned}\Theta(a, p, s, z, \phi_p, \phi_s, z', \phi'_p, \phi'_s) &= (1 - \pi_o) \mathbf{1}\{S(a, p, 0, z', \phi'_p, \phi'_s) \\ &\quad > S(a, p, s, z, \phi_p, \phi_s)\} \\ &\quad + \pi_o \mathbf{1}\{S(a, \kappa p, 0, z', \phi'_p, \phi'_s) - L_u(a, p) \\ &\quad > S(a, p, s, z, \phi_p, \phi_s)\}\end{aligned}$$

is the unconditional switching probability.

C.1. Stationary density of matches

Outflows from state $(a, p, s, z, \phi_p, \phi_s)$ have five sources: exogenous worker death and firm destruction, poaching events in which the worker switches jobs, productivity shocks that change z (whether the match continues at the new z or dissolves), and accepted learning opportunities that change (p, s) . Inflows have four sources: new matches formed from unemployed workers and vacancies, new matches formed from job-to-job transitions induced by poaching, productivity shocks at other matches in $(a, p, s, z', \phi_p, \phi_s)$ that land at z , and accepted learning opportunities at other matches in $(a, p', s', z, \phi_p, \phi_s)$ whose increment lands at (p, s) . Match formation always starts with zero firm-specific human capital, so the unemployment-to-employment and job-to-job inflows enter on the boundary $s = 0$ at whatever p the worker carries into the match (or κp for cross-occupation transitions); strictly positive $s > 0$ arises only through accepted learning opportunities within existing matches. Stationarity implies that $\Gamma(a, p, s, z, \phi_p, \phi_s)$ must satisfy, for any point with $s > 0$,

$$(\delta_w + \delta_e + \eta\alpha) \iiint \Theta(a, p, s, z, \phi_p, \phi_s, z', \phi'_p, \phi'_s) v(z', \phi'_p, \phi'_s) dz' d\phi'_p d\phi'_s$$

$$\begin{aligned}
& + \delta_z + \delta_l \int \mathbf{1}\{S(a, p + \Delta h, s + (1 - \Delta)h, z, \phi_p, \phi_s) \geq S(a, p, s, z, \phi_p, \phi_s)\} \\
& \times \gamma_h(h|a, p, s, \phi_p, \phi_s) dh \Gamma(a, p, s, z, \phi_p, \phi_s) \\
& = \delta_z \mathbf{1}\{S(a, p, s, z, \phi_p, \phi_s) \geq 0\} \mathbf{1}\{\Pi_v(z, \phi_p, \phi_s) \geq 0\} \gamma_z(z) \int \Gamma(a, p, s, z', \phi_p, \phi_s) dz' \\
& + \delta_l \iint \Gamma(a, p', s', z, \phi_p, \phi_s) \mathcal{K}(p, s|p', s', \phi_p, \phi_s) \\
& \times \mathbf{1}\{S(a, p, s, z, \phi_p, \phi_s) \geq S(a, p', s', z, \phi_p, \phi_s)\} dp' ds', \tag{A5}
\end{aligned}$$

where $\Delta = \Delta(\phi_p, \phi_s)$, and

$$\begin{aligned}
& \mathcal{K}(p, s|p', s', \phi_p, \phi_s) \\
& = \int \gamma_h(h|a, p', s', \phi_p, \phi_s) \delta_D(p - p' - \Delta h) \delta_D(s - s' - (1 - \Delta)h) dh
\end{aligned}$$

is the transition density induced by learning opportunities, with $\delta_D(\cdot)$ denoting the Dirac delta function.

At the boundary $s = 0$, the density $\Gamma(a, p, 0, z, \phi_p, \phi_s)$ is pinned down by inflows from match formation. New matches enter at $s = 0$ from two sources: unemployment-to-employment and job-to-job transitions. Each source contributes a same-occupation piece (preserving the worker's p) and a cross-occupation piece (rescaling $p \mapsto \kappa p$ at match formation, which by change of variables introduces a $1/\kappa$ Jacobian factor on the inflow density). Stationarity implies

$$\begin{aligned}
& \mathcal{O}(a, p, 0, z, \phi_p, \phi_s) \Gamma(a, p, 0, z, \phi_p, \phi_s) \\
& = \alpha v(z, \phi_p, \phi_s) \left[(1 - \pi_o) u(a, p) \mathbf{1}\{S(a, p, 0, z, \phi_p, \phi_s) \geq 0\} \right. \\
& \quad \left. + \pi_o \frac{1}{\kappa} u(a, p/\kappa) \mathbf{1}\{S(a, p, 0, z, \phi_p, \phi_s) - L_u(a, p/\kappa) \geq 0\} \right] \\
& \quad + \eta \alpha v(z, \phi_p, \phi_s) \left[(1 - \pi_o) \int \cdots \int \Gamma(a, p, s', z', \phi'_p, \phi'_s) \right. \\
& \quad \times \mathbf{1}\{S(a, p, 0, z, \phi_p, \phi_s) > S(a, p, s', z', \phi'_p, \phi'_s)\} ds' dz' d\phi'_p d\phi'_s \\
& \quad \left. + \pi_o \frac{1}{\kappa} \int \cdots \int \Gamma(a, p/\kappa, s', z', \phi'_p, \phi'_s) \right. \\
& \quad \times \mathbf{1}\{S(a, p, 0, z, \phi_p, \phi_s) - L_u(a, p/\kappa) > S(a, p/\kappa, s', z', \phi'_p, \phi'_s)\} ds' dz' d\phi'_p d\phi'_s \Big] \\
& \quad + \delta_z \mathbf{1}\{S(a, p, 0, z, \phi_p, \phi_s) \geq 0\} \mathbf{1}\{\Pi_v(z, \phi_p, \phi_s) \geq 0\} \gamma_z(z) \int \Gamma(a, p, 0, z', \phi_p, \phi_s) dz', \tag{A6}
\end{aligned}$$

where $\mathcal{O}(a, p, 0, z, \phi_p, \phi_s)$ is the total outflow rate from state $(a, p, 0, z, \phi_p, \phi_s)$, defined as the parenthesized prefactor of Γ in (A5) evaluated at $s = 0$. No learning inflow appears at the boundary: a learning opportunity at source (p', s') lands at the target $(p, 0)$ only if $(1 - \Delta)h = -s'$ and $p - p' = \Delta h$, which

(since $h \geq 0$, $s' \geq 0$, and $1 - \Delta \geq 0$) admits only the degenerate solution $h = s' = 0$, a measure-zero event under any continuous γ_h . The factors $\frac{1}{\kappa} u(a, p/\kappa)$ and $\frac{1}{\kappa} \Gamma(a, p/\kappa, s', z', \phi'_p, \phi'_s)$ in the cross-occupation inflows are to be read as the push-forwards of u and Γ under the rescaling $p_{\text{old}} \mapsto \kappa p_{\text{old}}$, and apply uniformly to the smooth-density and atomic components of those distributions. For a smooth source density, $\frac{1}{\kappa}$ is the change-of-variables Jacobian; for the Dirac atom of newborn entrants at $p_{\text{old}} = 0$ (and the analogous atom in Γ), the Dirac scaling identity $\delta_D(p/\kappa) = \kappa \delta_D(p)$ makes the $\frac{1}{\kappa}$ cancel against the κ from the delta, leaving the atom mass at $p = 0$ exactly preserved.

C.2. Stationary density of unemployed workers

Outflows from state (a, p) occur if the unemployed worker dies or gets matched with a vacancy. Inflows into state (a, p) have three sources. First, newborn entrants enter the economy at rate δ_w per unit time (replacing the deceased), draw a from γ_a , and start with $p = 0$; this inflow is therefore concentrated at the boundary $p = 0$. Second, a match involving a worker of type (a, p) may dissolve exogenously through firm destruction, sending the worker back to unemployment at her current p . Third, a productivity shock at the worker's match may drive surplus negative, dissolving the match endogenously and again returning the worker to unemployment at her current p . The single equation below covers both the smooth-density part of $u(a, p)$ at $p > 0$ and the newborn atom at $p = 0$ simultaneously. Stationarity implies that $u(a, p)$ satisfies

$$\begin{aligned} & \left(\delta_w + \alpha \iiint \Lambda(a, p, z', \phi'_p, \phi'_s) \nu(z', \phi'_p, \phi'_s) dz' d\phi'_p d\phi'_s \right) u(a, p) \\ &= \delta_w \gamma_a(a) \delta_D(p) + \delta_e \int \cdots \int \Gamma(a, p, s, z, \phi_p, \phi_s) ds dz d\phi_p d\phi_s \\ &+ \delta_z \int \cdots \int \Gamma(a, p, s, z, \phi_p, \phi_s) \\ &\quad \times \left[\int \mathbf{1}\{\min\{S(a, p, s, z', \phi_p, \phi_s), \Pi_\nu(z', \phi_p, \phi_s)\} < 0\} \gamma_z(z') dz' \right] ds dz d\phi_p d\phi_s, \end{aligned} \quad (\text{A7})$$

where the integral against $\Lambda \nu$ on the LHS is the per-worker matching rate—the vacancy meeting rate weighted by the unconditional match-formation probability Λ , integrated over vacancy types—and $\delta_D(p)$ is the Dirac delta at $p = 0$ encoding the assumption that all entrants start with zero portable human capital.

C.3. Stationary density of vacancies

Outflows from state (z, ϕ_p, ϕ_s) occur if the vacant firm is destroyed by an exogenous shock, if it successfully matches with a worker (employed or unemployed), or if it receives a productivity shock—which either moves the vacancy to a different productivity state or, if the new productivity is infeasible,

prompts endogenous exit. Inflows occur when an entrant firm draws (z, ϕ_p, ϕ_s) from $\gamma_z \gamma_\phi$, when a match featuring a type- (z, ϕ_p, ϕ_s) firm is dissolved exogenously by worker death, when such a match becomes infeasible after a productivity shock that lands in state (z, ϕ_p, ϕ_s) , when an existing vacancy at (z', ϕ_p, ϕ_s) receives a productivity shock that lands at z , or when a worker employed at a match featuring a type- (z, ϕ_p, ϕ_s) firm is poached. The equation is understood to hold on the feasible support $\{(z, \phi_p, \phi_s) : \Pi_v(z, \phi_p, \phi_s) \geq 0\}$, with $v \equiv 0$ off-support. Stationarity implies

$$\begin{aligned}
& \left[\delta_e + \alpha \iint u(a, p) \Lambda(a, p, z, \phi_p, \phi_s) da dp \right. \\
& \quad + \eta \alpha \int \cdots \int \Gamma(a, p, s, z', \phi'_p, \phi'_s) \Theta(a, p, s, z', \phi'_p, \phi'_s, z, \phi_p, \phi_s) \\
& \quad \quad \times da dp ds dz' d\phi'_p d\phi'_s \\
& \quad \left. + \delta_z \right] v(z, \phi_p, \phi_s) \\
& = \mu_J \gamma_z(z) \gamma_\phi(\phi_p, \phi_s) + \delta_w \iiint \Gamma(a, p, s, z, \phi_p, \phi_s) da dp ds \\
& \quad + \delta_z \gamma_z(z) \int \cdots \int \Gamma(a, p, s, z', \phi_p, \phi_s) \mathbf{1}\{S(a, p, s, z, \phi_p, \phi_s) < 0\} da dp ds dz' \\
& \quad + \delta_z \gamma_z(z) \int v(z', \phi_p, \phi_s) dz' \\
& \quad + \eta \alpha \int \cdots \int \Gamma(a, p, s, z, \phi_p, \phi_s) \Theta(a, p, s, z, \phi_p, \phi_s, z', \phi'_p, \phi'_s) \\
& \quad \quad \times v(z', \phi'_p, \phi'_s) da dp ds dz' d\phi'_p d\phi'_s.
\end{aligned} \tag{A8}$$

The matching outflow on the LHS uses the unconditional match-formation indicator Λ for unemployed workers and the unconditional switching indicator Θ for employed workers (defined above), with z' ranging over the worker's current firm in the job-to-job term. The total productivity-shock outflow is δ_z : every shock removes the vacancy from state (z, ϕ_p, ϕ_s) , with the exit option absorbed into the conditioning on the feasible support. In the productivity-shock inflows, z' is the pre-shock source state; this matches the source-as- z' convention used in (A5). In the JJ-poach inflow on the RHS, the convention is reversed: (z, ϕ_p, ϕ_s) is the worker's pre-poach (and the equation's) cell, while (z', ϕ'_p, ϕ'_s) is the destination vacancy the worker accepts.

The entry rate μ_J is the rate at which potential entrants arrive and draw a type from $\gamma_z \gamma_\phi$; only those with $\Pi_v \geq 0$ post a vacancy, the rest exit immediately. Stationarity of J requires the rate at which entrants post a vacancy, μ_J times the probability that a drawn type has $\Pi_v \geq 0$, to equal the total rate at which firms leave the system:

$$\mu_J \iiint \mathbf{1}\{\Pi_v(z, \phi_p, \phi_s) \geq 0\} \gamma_z(z) \gamma_\phi(\phi_p, \phi_s) dz d\phi_p d\phi_s$$

$$\begin{aligned}
&= \delta_e J + \delta_z \iiint v(z, \phi_p, \phi_s) \int \mathbf{1}\{\Pi_v(z', \phi_p, \phi_s) < 0\} \gamma_z(z') dz' dz d\phi_p d\phi_s \\
&+ \delta_z \int \dots \int \Gamma(a, p, s, z, \phi_p, \phi_s) \int \mathbf{1}\{\Pi_v(z', \phi_p, \phi_s) < 0\} \gamma_z(z') \\
&\quad \times dz' da dp ds dz d\phi_p d\phi_s.
\end{aligned} \tag{A9}$$

The first term is exogenous firm destruction. The second is endogenous exit from vacancies whose productivity is shocked into the infeasible region. The third is endogenous exit from filled firms whose post-shock vacancy value is negative (recall that a match is destroyed if the firm's continuation value as a vacancy at the new productivity is negative, meaning that upon match dissolution, the firm exits rather than becoming vacant).

Appendix D. Solution Algorithm

A stationary equilibrium consists of three value functions, $S(a, p, s, z, \phi_p, \phi_s)$, $W_u(a, p)$, and $\Pi_v(z, \phi_p, \phi_s)$; three densities, $\Gamma(a, p, s, z, \phi_p, \phi_s)$, $u(a, p)$, and $v(z, \phi_p, \phi_s)$; and the meeting rate α . The algorithm takes α and μ_j as given and solves for $(S, W_u, \Pi_v, \Gamma, u, v)$. The matching efficiency A and the firm entry cost c_e are recovered ex post from the converged objects using (3) and the free-entry condition (20) holding with equality, respectively.

The algorithm is a nested fixed point with an outer Gauss–Seidel loop that alternates between updating (S, W_u, Π_v) given (Γ, u, v) in Step 1, and updating (Γ, u, v) given (S, W_u, Π_v) in Step 2.

Throughout, let $N_a, N_p, N_s, N_z, N_{\phi_p}, N_{\phi_s}$ denote the number of grid points in state variables $a, p, s, z, \phi_p, \phi_s$ respectively. Define the surplus grid size $N_S = N_a N_p N_s N_z N_{\phi_p} N_{\phi_s}$, the firm grid size $N_F = N_z N_{\phi_p} N_{\phi_s}$, and the unemployed grid size $N_u = N_a N_p$.

Step 0 — Discretization and initialization

Grids and quadrature. Build a grid on each state variable $a, p, s, z, \phi_p, \phi_s$. The grids for p and s include $p = 0$ and $s = 0$ as boundary cells and are power-spaced, $x_i = x_{\max}((i-1)/(N-1))^c$ with $c > 1$, so that nodes are densest near zero, where new workers and new matches start. Discretize the productivity distribution γ_z on the z grid via Rouwenhorst. Discretize the worker characteristic a via Gauss–Jacobi quadrature on $[-1, 1]$ with parameters $(a_\beta - 1, a_\alpha - 1)$, mapped to $[0, 1]$ by $a = (x + 1)/2$. Pre-compute Gauss–Laguerre nodes and weights for the learning kernel γ_h , and Gauss–Hermite nodes and weights for the entrants' draws from γ_z and γ_ϕ .

Initialization. Set $W_u^{(0)} \equiv b/(r + \delta_w)$; $S^{(0)} \equiv 0$; $\Pi_v^{(0)} \equiv 0$ on the full grid (so the initial vacancy support is the full firm grid); $u^{(0)}(a, p) = \gamma_a(a) \delta_D(p)$ as a boundary atom; $v^{(0)}(z, \phi_p, \phi_s) \propto \gamma_z(z) \gamma_\phi(\phi_p, \phi_s)$

on the full grid; $\Gamma^{(0)} \equiv 0$. Set the six per-unknown outer tolerances ε_S , ε_{W_u} , ε_{Π_v} , ε_Γ , ε_u , and ε_v ; the inner-tolerance tightening factor $\rho \geq 1$ (so that Steps 1 and 2 converge to ε_x/ρ internally); the inner and outer damping factors $\lambda_{\text{in}}, \lambda_{\text{out}} \in (0, 1]$; the inner and outer iteration caps $I_{\text{in}}, I_{\text{out}} \in \mathbb{N}$; and the smoothing scale $\varepsilon_\sigma > 0$.

Step 1 — Value update given densities

Inputs: $\Gamma^{(n)}$, $u^{(n)}$, and $v^{(n)}$.

Outputs: $W_u^{(n+1)}$, $S^{(n+1)}$, and $\Pi_v^{(n+1)}$.

The value functions are updated by damped Picard iteration on their fixed-point equations (15), (19), (16). The Picard operator $\mathcal{T}_x$ for each equation is the RHS divided by the LHS coefficient, evaluated at the current iterate; for (19) and (16) the productivity-shock and learning-acceptance self-terms ($-\delta_z S$, $-\delta_l \mathbf{1}\{S^{\text{new}} \geq S\} \cdot S$, and $-\delta_z \Pi_v$) are first absorbed into the LHS coefficient. In (19), the hard learning-acceptance indicator is replaced by the sigmoid

$$\sigma_{\varepsilon_\sigma}(S^{\text{new}} - S) = \frac{1}{1 + \exp(-(S^{\text{new}} - S)/\varepsilon_\sigma)},$$

with smoothing scale $\varepsilon_\sigma > 0$; the original equation is recovered as $\varepsilon_\sigma \rightarrow 0$.

The three substeps—1a updating W_u , 1b updating S given the freshly updated W_u , 1c updating Π_v given the freshly updated W_u, S —are executed as a Gauss–Seidel sweep $1a \rightarrow 1b \rightarrow 1c$ with inner damping factor λ_{in} . The sweep is repeated up to I_{in} times, until the pre-damping residuals $r_x = \|\mathcal{T}_x(x) - x\|_\infty$ for $x \in \{W_u, S, \Pi_v\}$ are each below the inner tolerances ε_x/ρ . The factor ρ tightens the inner stopping rule relative to the outer tolerances of Step 3, so inner residuals cannot contaminate the outer convergence check.

Step 2 — Density update given values

Inputs: $W_u^{(n+1)}$, $S^{(n+1)}$, $\Pi_v^{(n+1)}$, $\Gamma^{(n)}$, $u^{(n)}$, and $v^{(n)}$.

Outputs: $\Gamma^{(n+1)}$, $u^{(n+1)}$, and $v^{(n+1)}$.

The densities are updated by damped Picard iteration on their stationary KFEs (A5)–(A8). Each Picard operator $\mathcal{T}_x$ is the RHS of the equation divided by its LHS coefficient at the current iterate, with all $-x$ -self-terms first absorbed into the LHS. The hard match-acceptance and switching indicators inside Λ , Θ , Ψ , and the $\mathbf{1}\{S \geq 0\}$ flags in the productivity-shock channel are evaluated at the converged $S^{(n+1)}$ and $\Pi_v^{(n+1)}$, and held fixed throughout the Step 2 sub-iteration.

Atoms and smooth densities. Match formation places newborn workers at $p = 0$ in u and new matches at $s = 0$ in Γ . Each atom is carried as a separate object alongside the smooth interior density: a scalar per a -cell for the u -atom at $p = 0$, and an array indexed by $(a, p, z, \phi_p, \phi_s)$ for the Γ -atom at $s = 0$. Every aggregate integral against u or Γ is then split into a smooth quadrature and an atom evaluation. For the cross-occupation rescaling $p \mapsto \kappa p$ the source u -atom at $p = 0$ maps to the target at $p = 0$, with the smooth Jacobian $1/\kappa$ and the Dirac scaling identity $\delta_D(p/\kappa) = \kappa \delta_D(p)$ canceling exactly, so atom mass at $p = 0$ is preserved.

Support masking. The vacancy density satisfies $v \equiv 0$ on $\{(z, \phi_p, \phi_s) : \Pi_v(z, \phi_p, \phi_s) < 0\}$, and the match density satisfies $\Gamma \equiv 0$ on $\{(a, p, s, z, \phi_p, \phi_s) : S < 0\} \cup \{(z, \phi_p, \phi_s) : \Pi_v < 0\}$. These supports are imposed explicitly at every Picard sweep by zeroing the corresponding cells (and the corresponding Γ_{atom} entries) after the divide-by-LHS update; the indicators inside the integrands of (A5)–(A8) ensure that the inflow contributions from infeasible source cells are also discarded.

The three substeps—2a updating v from (A8), 2b updating u from (A7) given the freshly updated v , 2c updating Γ from (A5)–(A6) given the freshly updated u, v —are executed as a Gauss–Seidel sweep $2a \rightarrow 2b \rightarrow 2c$ with inner damping factor λ_{in} . Within 2b the smooth piece $u(a, p)$ at $p > 0$ and the atom $u_{\text{atom}}(a)$ at $p = 0$ are updated jointly: only the atom carries the boundary source $\delta_w \gamma_a(a)$ from (A7), while the firm-destruction and productivity-shock dissolution inflows on the RHS, which are smooth in p , populate $u(a, p)$ at all grid points (including the boundary cell $p = 0$ as the smooth limit). Within 2c the interior equation (A5) updates Γ for $s > 0$ and the boundary equation (A6) updates Γ_{atom} at $s = 0$; the cross-occupation push-forward $p \mapsto \kappa p$ in the boundary equation is implemented as a mass-conserving deposition: each source cell’s mass at p_{src} is distributed into the two target cells bracketing κp_{src} via 2-point cloud-in-cell linear weights $D[i_{\text{src}}, i_{\text{tgt}}]$, with $\sum_{i_{\text{tgt}}} D[i_{\text{src}}, i_{\text{tgt}}] = 1$ at every interior source so total mass is preserved exactly on the discrete grid. In the continuum limit this recovers the $(1/\kappa) u(a, p/\kappa)$ Jacobian factor in (A6). The atom at $p_{\text{src}} = 0$ maps to $p_{\text{tgt}} = 0$ with $D[1, 1] = 1$, consistent with the Dirac scaling identity $\delta_D(p/\kappa) = \kappa \delta_D(p)$. The sweep is repeated up to I_{in} times, until the pre-damping residuals $r_x = \|\mathcal{T}_x(x) - x\|_\infty$ for $x \in \{v, u, \Gamma\}$ are each below the inner tolerances ε_x/ρ .

Total worker mass. In the continuum the worker conservation identity

$$\int u \, da \, dp + \int \Gamma \, da \, dp \, ds \, dz \, d\phi_p \, d\phi_s = 1,$$

is a corollary of the death–replacement structure: workers exit at rate δ_w uniformly across (u, Γ) and are replaced at rate $\delta_w \gamma_a(a) \delta_D(p)$ with $\int \gamma_a(a) \, da = 1$, so total worker mass is invariant under the KFEs. On the discrete grid this identity is broken by quadrature error in the matching integrals; the cross-occupation push-forward is mass-conserving by construction and does not contribute to the drift. After Step 2 converges, the total worker mass $M = \int u + \int \Gamma$ is computed from the discrete sums, including both the smooth interior and the atoms, and is then used to rescale $u^{(n+1)}, u_{\text{atom}}^{(n+1)}, \Gamma^{(n+1)}, \Gamma_{\text{atom}}^{(n+1)}$ to enforce the identity exactly. The rescaling is a numerical safety net for a continuum-level conservation

law, not an additional equilibrium condition; the firm aggregate $J = F + V$ is determined endogenously by μ_J and the exit rates and is not separately normalized.

Step 3 — Outer convergence

Inputs: $\Gamma^{(n)}, u^{(n)}, v^{(n)}, W_u^{(n)}, S^{(n)}$, and $\Pi_v^{(n)}$.

Outputs: $\Gamma^{(n+1)}, u^{(n+1)}, v^{(n+1)}, W_u^{(n+1)}, S^{(n+1)}$, and $\Pi_v^{(n+1)}$.

Step 3 is a damped Gauss–Seidel iteration that wraps Step 1 and Step 2. To assess convergence, the six per-unknown sup-norm residuals

$$\Delta_x = \|x^{\text{new}} - x^{\text{snap}}\|_{\infty}, \quad x \in \{S, W_u, \Pi_v, \Gamma, u, v\},$$

are computed, the residuals for Γ and u taking the maximum across the smooth and atom branches. Convergence is achieved when $\Delta_x < \varepsilon_x$ for all six unknowns simultaneously.

Step 4 — Recover c_e and A

Inputs: Converged Π_v, u, Γ, v from Step 3.

Outputs: Firm entry cost c_e , and matching efficiency A .

The firm entry cost is recovered from the free-entry condition (20). The aggregate stocks U, F, V are obtained by summing the converged densities across smooth and atom branches. Lastly, the matching efficiency is recovered from the matching function (3). In the counterfactual experiment of Section 6, the matching efficiency A and the entry cost c_e are held at their estimated values, so the meeting rate α and the entry rate μ_J are solved for: Steps 0 to 3 are repeated at the new parameters and at candidate values of (α, μ_J) until the A and c_e recovered in this step equal their estimated values.

Appendix E. Identification of the Remaining Parameters

Section 4.2 explains which moments identify the productivity process, the learning technology and the distribution of usage intensities, and Section 4 how the moments are measured on the two sides. This appendix gives the corresponding heuristic for the remaining parameters, in the order of Table 2.

Search. With the elasticity μ fixed, the matching process leaves two parameters to identify: the matching efficiency A and the relative search intensity of employed workers η . The parameter A is identified from the unemployment-to-employment transition rate, as the job-finding hazard of unemployed workers depends only on the matching technology. The parameter η is identified from the job-to-job

transition rate: once A is pinned down, the remaining variation in job-to-job mobility is driven by the search intensity of employed workers.

Costs. The firm entry cost c_e is identified from the firm-to-worker ratio: the equilibrium mass of firms adjusts until entry is marginally profitable. The vacancy-posting cost c_v is identified from the vacancy-to-firm ratio. Free entry pins down the value of a vacancy, while the flow cost of maintaining an open position determines the equilibrium vacancy margin.

Bargaining. The worker bargaining power β affects the *division* of match output between the worker and the firm, but not its *size*. The worker's share of output provides a direct measure of this division. In the data, a worker's share is her hourly wage times 1,820 hours, a full-time year, over her firm's value added per worker, both adjusted as in Section 2.5; in the model, it is her wage over match output, scaled by 0.65, one minus the capital share of value added, because value added in the data also remunerates capital. Shares are winsorized at the 1st and 99th percentiles. Higher bargaining power grants the worker a larger fraction of every unit of output and shifts the whole distribution of shares upward. Therefore, the mean of the worker's share identifies β .

Home production. The home-production flow b affects the value of unemployment and therefore workers' selectivity. A higher b raises the surplus that a meeting with an unemployed worker must generate for a match to form. Offers that would have paid the lowest starting wages are then turned down, which thins the lower tail of the starting wages of hires out of unemployment. A match's starting wage is the wage of its first full-time spell-year on a standard contract, observed in its first year, and a hire out of unemployment is a new match that directly follows an unemployment spell (Appendix A). The gap between the median and the 10th percentile of these starting wages identifies b .

Separation. In the model, a match can end in unemployment for three reasons: the firm is destroyed by an exogenous shock, which arrives at rate δ_e ; a sufficiently adverse productivity shock turns the match surplus negative and the match is dissolved; or the shock makes the firm unprofitable even as a vacancy and the firm exits. The employment-to-unemployment rate sums these causes, so on its own it cannot identify δ_e . The productivity process separates them. It is pinned down by the output moments (Section 4.2), which are unaffected by δ_e : an exogenous destruction shock removes a match, but it does not change the output of the matches that survive. Given the productivity process, the model implies the frequency of endogenous separations. The part of the employment-to-unemployment rate that endogenous separations cannot account for identifies δ_e .

Workers leave the labor market at rate δ_w . This rate sets how long careers last and therefore how much experience the employed have accumulated, so mean experience identifies δ_w .

Occupational shocks. Two parameters govern occupational shocks that induce portable human capital losses when new matches are formed. The share of cross-occupation meetings is controlled by π_o , while

the reduction in the amount of portable human capital a worker retains after the move is determined by κ .

The rate of cross-occupation hires out of unemployment provides the cleanest identification margin for π_o . The unemployed tend to carry little portable capital: entrants begin with none, while workers with substantial portable capital tend to find jobs quickly and are less likely to remain unemployed. As a result, the unemployment pool is concentrated among workers for whom the loss of portable capital from switching occupations is relatively small. For these workers, a cross-occupation offer is therefore almost equivalent to a same-occupation offer, making acceptance decisions largely independent of κ . The observed rate at which unemployed workers are re-employed in a different occupation consequently reflects the underlying composition of offers and identifies π_o with limited influence from switching costs.

The retention factor κ is identified from the wage loss of displaced workers who are re-employed in a different occupation, measured two years after the closure and net of the wage growth of stayers. A firm closure is an involuntary move: the worker did not choose to leave, so her move is not selected on the value of switching. When she is re-employed in a different occupation, she keeps only a fraction κ of her portable stock, and her wage change reflects that loss.

Pre-market skill. The Beta shape parameters (a_α, a_β) govern the spread and the asymmetry of the entry distribution of pre-market skill a . They are identified from the distribution of wages at the start of careers, before human capital accumulation and job-ladder dynamics distort the wage distribution, so entry wages primarily reflect the underlying distribution of pre-market skill. The variance of log wages among workers with less than one year of experience captures the spread of the entry distribution, while their asymmetry captures its skew.

Appendix F. The Model without Firm-Specific Capital

Section 5 confronts the estimated model and a restricted model without firm-specific human capital with the same untargeted facts. This appendix describes how the restricted model is built and estimated, and how it fits the targeted moments that inform the distribution of firm-specific usage intensities in the estimated model.

Construction. In the restricted model, no job uses or builds firm-specific human capital. Every learning increment is allocated to portable human capital: the allocation share is set to $\Delta = 1$ in every job, equivalently $\phi_s = 0$, so the firm-specific stock stays at $s = 0$ over every career and the intensity ϕ_s plays no role in output, learning or wages. The parameters of its distribution, μ_s , σ_s and ρ_ϕ , therefore drop out. Everything else is as in Section 3: the job ladder, with sequential auctions between the current employer and the poacher; bargaining; firm productivity shocks; cross-occupation meetings, with

TABLE A1. Fit on the targets of firm-specific capital: data, estimated model and restricted model.

	Data	Full	Restricted
Displacement loss, incumbents – newcomers	–0.087	–0.079	–0.035
Var. of displacement losses, incumbents – newcomers	0.035	0.033	0.012
Slope of loss on starting wage, experience 10+ – 0–1	0.135	0.118	0.068

NOTES: The three targeted moments of the μ_s , σ_s and ρ_ϕ rows of Table 2, measured as there. Full: the model of Section 3 at the estimates of Table 2. Restricted: the model without firm-specific human capital, re-estimated on the same targeted moments. Displacement losses are in log units and net of the wage growth of stayers.

the haircut κ on portable human capital; and the heterogeneity across jobs in the usage intensity of portable human capital, ϕ_p . The restricted model has 19 estimated parameters, against 22 in the estimated model. It is estimated by SMM on the same full vector of targeted moments and with the same diagonal weighting matrix (Section 4), so it is over-identified, and its simulated panel is built and sampled exactly as the estimated model's.

Objective. At its estimates, the restricted model's SMM objective is 118.3, against 41.7 for the estimated model. With a diagonal weighting matrix this comparison is descriptive, not a test.

Fit on the targets of firm-specific capital. Table A1 reports both models' fit on the three targeted moments that, in the estimated model, inform μ_s , σ_s and ρ_ϕ (Section 4.2). In the restricted model, displaced incumbents lose more than newcomers only through the job-ladder rents they had accumulated and through the haircut on portable human capital when they move to a new occupation, and their losses differ from one another only through these channels and the dispersion of ϕ_p . The restricted model's gap between incumbents' and newcomers' mean losses is –0.035, against –0.087 in the data; its variance gap is 0.012, against 0.035; and the change of the loss-on-starting-wage slope across experience at hire is 0.068, against 0.135.

Appendix G. Fact 1: Measuring the Incumbent Premiums

Section 5.1 reads the incumbent premiums off the slope of a firm's output per worker and of its mean log wage on its newcomer share. This appendix derives that slope from the accounting identity, bounds the approximation it uses, and states the condition under which the regressions of Table A2 recover the premiums. Firm and year subscripts are dropped where no change over time is involved.

Output. Output per worker is the share-weighted average of what incumbents and newcomers produce, $y = (1 - n) y^I + n y^N$. Factoring out the incumbents' output,

$$y = y^I \left[1 - n \left(1 - \frac{y^N}{y^I} \right) \right],$$

and with $P_y = \log y^I - \log y^N$, so that $y^N/y^I = e^{-P_y}$,

$$\log y = \log y^I + \log [1 - n (1 - e^{-P_y})]. \quad (\text{A10})$$

This is exact. The second term is the composition discount: it is zero when the firm has no newcomers or when the two groups produce the same, and it falls as either the newcomer share or the premium rises. As a function of P_y it equals zero at $P_y = 0$ and has derivative $-n$ there, so to first order in the premium $\log y \approx \log y^I - P_y n$, which is equation (23). The exact slope of (A10) with respect to n is $-(1 - e^{-P_y})/[1 - n(1 - e^{-P_y})]$. Its gap to $-P_y$ is of second order in the premium: at the premium of Table A2 it is below half a standard error at every newcomer share. Because value added is observed only as a firm total, y^I and y^N are averages of levels, and the output premium is the log of their ratio.

Wages. Equation (24) holds exactly because the mean log wage is the full-time-equivalent-weighted mean of the workers' log wages and the newcomer share uses the same weights: splitting a weighted mean between two groups of workers is an identity, with no step in levels. The wage premium P_w is the gap between the two groups' mean log wages.

From the identities to the regressions. Equation (23) holds firm by firm and year by year, with $\log y_{ft}^I$ unobserved. Its change between consecutive years is $\Delta \log y_{ft} \approx \Delta \log y_{ft}^I - \Delta(P_{y,ft} n_{ft})$. Firm levels drop out: a productive firm with low turnover has a high y_{ft}^I and a low n_{ft} in every year, which would confound a cross-section of levels but not the firm's own changes. Write the premium as a firm component and a deviation, $P_{y,ft} = P_{y,f} + p_{ft}$. Then $\Delta(P_{y,ft} n_{ft}) = P_{y,f} \Delta n_{ft} + \Delta(p_{ft} n_{ft})$: the first term is the firm's premium times the change in its share, and the second is nonzero only when the premium moves between the two years; it joins the residual. Writing the change in incumbents' output as an industry-year effect, a linear term in the controls, and a residual, $\Delta \log y_{ft}^I = d_{y,qt} + \gamma'_y \Delta x_{ft} + \varepsilon_{y,ft}$, gives (26) with $P_{y,f}$ in place of P_y ; the wage equation (27) follows in the same way from (24). The regression imposes a single P_y . Least squares then returns a weighted average of the firms' premiums, with weights proportional to the squared change in the newcomer share: up to the adjustment for the controls, the coefficient on Δn_{ft} converges to minus the sum of $P_{y,f} (\Delta n_{ft})^2$ over the sample divided by the sum of $(\Delta n_{ft})^2$. This average, and P_w likewise, is recovered if the residual, including the term $\Delta(p_{ft} n_{ft})$, is uncorrelated with Δn_{ft} given the controls and the industry-year effects: a change in the newcomer share must not coincide, for other reasons, with a change in what the firm's incumbents produce or earn. If firms hire when their incumbents' output rises, as in a good year for the firm, the covariance is positive, the coefficient on Δn_{ft} is less negative than $-P_y$, and the output premium

TABLE A2. Fact 1: incumbents and newcomers.

Incumbent over newcomer	Data	Full	Restricted
Output premium	5.6 (0.4)	5.6	4.1
Wage premium	4.3 (0.1)	4.0	3.7
Wedge	1.3 (0.4)	1.6	0.4

NOTES: Log points. Each row is minus the coefficient on the change in the firm's newcomer share in a regression of the change in its log value added per full-time-equivalent worker (output premium), in its mean log hourly wage (wage premium), or in the difference between the two (wedge). Controls: the change in mean age and in log tangible fixed assets per full-time-equivalent worker, and a dummy for each two-digit industry and year. Sample: 784,751 firm-year pairs, 2010–2011 to 2015–2016 (Appendix A). Standard errors clustered by firm in parentheses. Full and Restricted: the gaps between incumbents and newcomers at the same firm state and with the same potential experience, computed directly in the panels simulated from the estimated model and from the restricted model (Appendix G).

is understated; the same holds for the wage premium, and the wedge is understated if incumbents' wages respond less than their output to such a year. The wedge regression replaces the outcome by $\Delta(\log y_{ft} - \bar{w}_{ft})$ and keeps the regressors, so by linearity of least squares its coefficient is exactly $-(P_y - P_w)$ and its standard error is that of the difference.

The model counterpart. In the model, the premiums are computed directly. The unit is an employed worker-year of the simulated panel, classified as in the data: an incumbent if the worker held an employment spell at the same firm in the previous year, a newcomer otherwise. Each worker-year records the log output of the match and the log hourly wage, as hours-weighted means over the year. Worker-years are grouped into cells defined by the state of the employing firm—its productivity z and the usage intensities ϕ_p and ϕ_s of its job—and by potential experience, the years since entry into the labor market, which takes the place of age. Within each cell, the output premium is the gap between the mean log output of incumbents and that of newcomers, and the wage premium is the same gap in log wages. The reported premiums average the within-cell gaps over cells that contain both groups, weighting each cell by its employed time, and the wedge is their difference. Firms in the same state are identical in the model, so the comparison within a cell is the counterpart of the data's comparison within a firm, and no regression is needed to remove differences across firms.

The data's regressions would not measure the same object on the model's firms, each of which employs one worker at a time. Run on such a firm observed in two consecutive years, the change in the newcomer share is close to 0 or ± 1 . A change of -1 compares the same worker in her hiring year and the following year, so it measures her own growth rather than a gap between incumbents and newcomers; a change of $+1$ compares a newcomer with the predecessor she replaces. And because the firm hires only after its worker leaves, which happens more often after a fall in productivity, the arrival of a newcomer coincides with adverse productivity shocks, which would tend to bias the estimated wedge upward, the opposite of the bias in the data.

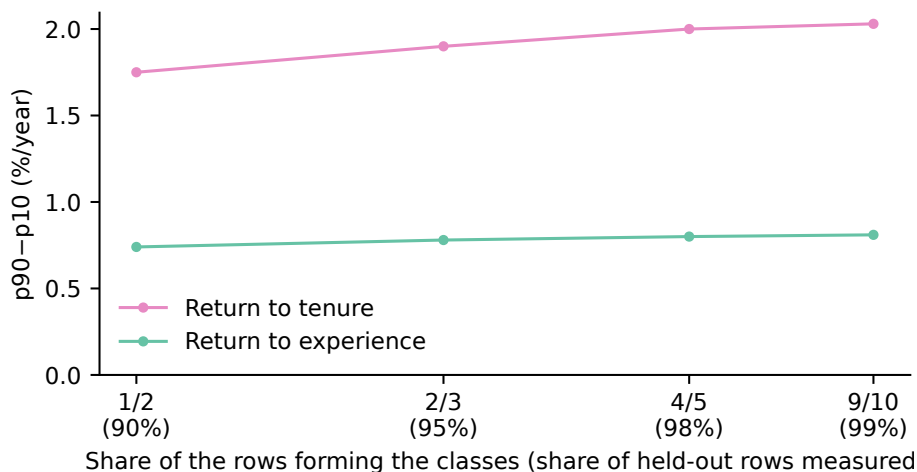


FIGURE A1. The dispersion of returns at G^* against the share of the sample forming the classes.

NOTES: Gap between the ninetieth and tenth percentiles of the job-class returns, in percent of the wage per year, weighted by spell-years, with $K = 5$ worker classes and $G^* = 40$ job classes, in the data. The spell-years are split at random into two, three, five or ten parts; the classes are formed on all parts but one and the returns estimated on the held-out part, every part held out once, and the estimates averaged. In parentheses, the share of held-out spell-years whose worker and job appear in the classifying parts.

Appendix H. The Grouped Fixed-Effects Estimator: Resolution and Sample

The dispersion of returns across job classes measured in Section 5 is read at the number of classes where it levels off, $G^* = 40$. This appendix reports two checks on what that plateau measures.

Placement. A job is placed in a class on the strength of its own spell-years in the parts that form the classes. The fewer those rows, the more often a job is placed in the wrong class, and the more the classes mix jobs with different returns, which pulls the class returns toward each other. The plateau could therefore reflect the limit of what the rows per job can resolve rather than the differences between jobs. Figure A1 repeats the estimation at G^* with the classes formed on one half, two thirds, four fifths and nine tenths of the spell-years, the returns estimated on the remaining rows and every part held out in turn. Both spreads rise from one half to two thirds and are flat from two thirds on, so the plateau is not an artifact of how much of each job's history the placement relies on. The share of held-out rows that can be measured, those whose worker and job appear in the classifying parts, is reported under each point; it rises with the classifying share and exceeds ninety-five percent from two thirds on, so the four points describe nearly the same jobs.

Sample. The model is validated on a panel simulated from it and put through the same estimator, so the two sides are measured at the same resolution only if a job is observed on comparable material in both. Table A3 compares the structure of the two samples: the simulated panel is sampled as the data is, one worker in twenty-five over nine years, and cut to the data's number of spell-years. A job in the

TABLE A3. The structure of the sample: data and simulated panel.

	Data	Model
Jobs	210,000	230,000
Workers per job, mean	2.1	1.9
Workers per job, ninetieth percentile	4	4
Spell-years per worker, mean	5.6	5.5
Spell-years per job, mean	5.2	4.9
Share of spell-years in jobs with one worker	45%	48%

NOTES: The employed spell-years on which the grouped fixed-effects estimator is fitted. Data: the DADS panel of Section 2, one worker in twenty-five. Model: the panel simulated from the estimated model, sampled in the same way and cut to the data's number of spell-years. A job is a firm–occupation cell in the data and a position in the model; workers per job counts the distinct workers observed in the job.

data is a firm–occupation cell; in the model it is a position at a firm. On both sides a job is observed through a small number of workers, most jobs through one, and a worker through about five and a half spell-years.