Endogenous Wage Rigidity and Layoffs

Andrei Zaloilo*†

Toulouse School of Economics

September 14, 2025

Click here for the latest version

Abstract

Firms rarely cut wages but often lay off workers. Using matched employer–employee data from France, I show that this pattern is not only widespread across firms but also systematic within firms across worker tenure: junior workers experience up to 5 times the layoff risk of senior workers, but suffer little to no wage cuts. Standard explanations of wage rigidity based on exogenous constraints cannot account for this evidence. I develop an equilibrium search model where firms employ risk-averse workers of varying match quality on dynamic contracts. Although theoretically the firm has to manage individual contracts with a continuum of employees, I show that the model can be solved tractably with tenure-specific contracts. Within a cohort, the firm is not allowed to wage discriminate, but can fire freely. This constraint induces firms to shed low-quality matches while keeping surviving matches within the cohort at unchanged wages. Calibrated to the French data, the model replicates the observed link between layoffs and wage rigidity across both firms and tenure, offering a new explanation for why firms fire rather than cut pay.

Keywords: productivity, layoffs, long-term contracts

JEL codes: D86, H23, J24, J31, J41, J62

*Email: andrei.zaloilo@tse-fr.eu

†I am extremely grateful to Christian Hellwig, Eugenia Gonzalez-Aguado, Dirk Krueger, and Nicolas Werquin for their invaluable advice and continuous guidance. I also thank Charles Brendon, Fabrice Collard, Alessandro Dovis, Patrick Feve, Alexandre Gaillard, Gerard Maideu Morera, Iourii Manovskii, Guillermo Ordonez, Eduardo Schaal, Simon Scheidegger, Philipp Wangner, Wenxuan Xu, Miguel Zerecero, and participants at various seminars for constructive comments and insightful discussions.

1 Introduction

Firms that lay off the most workers tend to cut wages the least. Why do some firms lay off workers rather than cut wages? In most of the literature, the answer is that firms cannot cut wages for exogenous reasons and therefore resort to layoffs. The assumption of exogenous wage rigidity—potentially justified by minimum wages, morale costs of cuts, or sectoral bargaining—is pervasive in macroeconomics. It is used to amplify employment fluctuations and firms' responses to monetary shocks. However, recent survey evidence (Davis and Krolikowski (2025), Bennedsen et al. (2025)) suggests that firms often have discretion and sometimes actively prefer layoffs to feasible wage cuts.

In this paper, I offer an alternative explanation for why firms lay off workers instead of cutting wages. Rather than lay off because they must, firms *choose* to lay off and then avoid cutting survivors' wages. For intuition, consider a survey question in Bennedsen et al. (2025): "Why didn't you lower pay instead of laying off employees?" The modal answer is: "Layoffs give better control over who leaves the firm." I interpret this as firms finding some workers more productive than others and preferring layoffs as a tool to remove lower-productivity matches (whereas wage cuts could induce valuable workers to exit). I build a dynamic model with a frictional labor market in which both unemployed and employed workers search. Each period, firms hire workers; each match has a quality realized after work begins and observed only by the firm. To facilitate that layoffs are the instrument for culling poor matches, the firm cannot wage-discriminate within a hiring cohort (workers hired in the same period). Layoffs are thus the only tool to affect a cohort's composition. When used, they raise the cohort's average quality and reduce the firm's incentive to cut survivors' wages.

Beyond surveys, I use French matched employer–employee data from 2009–2019 to discipline and validate the model. First, I confirm the literature's finding (Ehrlich and Montes (2024)) that firms which lay off the most exhibit the most rigid wages. I document the layoff rate and the pass-through of firm-level productivity shocks to the wages of job stayers. Grouping firms by average layoff rates, I find that after a positive productivity shock normalized to 100%, wages in firms firing the least increase by 1.7%, whereas wages in firms firing the most change by less than 1%.

Furthermore, the connection between layoffs and wage rigidity holds not only across firms but also within firms, across tenure. Workers with tenure below two years face layoff rates of 5–11%, compared with 2% for workers with four to five years of tenure. By contrast, wages of workers with less than one year of tenure essentially do not respond to negative productivity shocks, whereas more senior workers' wages fall by 2.4% in response to a 100% productivity shock.

Lastly, I examine relative wage changes between junior and senior workers when a firm begins layoffs. Whenever firms lay off workers, a wage gap opens between senior and junior contracts.

To understand these facts, the paper builds an equilibrium model of the labor market with search frictions and one-sided limited commitment on the worker side. Firms employ a continuum of workers, exhibit decreasing returns to scale, and face firm-level productivity shocks. Search is directed: firms post dynamic contracts to attract workers. Contracts specify future productivity-history-contingent wages and layoff risk. Workers are risk-averse and cannot commit to stay; firms can credibly insure workers against future productivity shocks. Firms thus face a trade-off between insuring workers against future losses and incentivizing workers to stay (leave) when they are most (least) productive. While ex-ante homogeneous, all matches receive an initial permanent shock at the start of employment. Firm-level productivity is common knowledge; the match-specific shock is observed only by the firm.

Because firms have decreasing returns to scale, they manage contracts jointly across employees. In principle, this renders the model intractable: in a recursive formulation, the firm would need to track contracts for a continuum of workers, producing an infinite-dimensional state space. Fortunately, with directed search, within a given period a firm hires all workers on the same contract. It is therefore sufficient to track contracts by cohort. This discretizes the state space and makes it finite for firms of finite age: a firm of age ten employs at most ten cohorts. Moreover, up to an approximation, the state space can be bounded: I show theoretically and empirically that cohort wages tend to converge. I use this to justify pooling workers beyond a tenure threshold onto the same contract. Thus, instead of optimizing a continuum of contracts, it suffices to work with a finite set of tenure-specific contracts.

I assume firms cannot wage-discriminate by match quality within a cohort. Layoffs are therefore the only way to affect within-cohort quality. In low-productivity states, layoffs serve two purposes: they shed labor that is overpaid relative to productivity and cleanse the workforce of poor matches. In cohorts that experience layoffs, surviving workers are of higher average quality. The firm is then more inclined to retain survivors and cuts their wages less than in cohorts whose composition did not change. This mechanism is central to the paper's view of layoffs and pay cuts: wherever the firm chooses to fire workers, it has less desire to cut survivors' wages.

Unlike models with exogenous wage rigidity, this framework explains the layoff–wage connection both between and within firms. When choosing whom to lay off, it is optimal to eliminate poor matches—typically the least costly workers to dismiss, often juniors. The model thus generates relatively higher layoff rates among the most junior workers. Individual layoff risk depends on both absolute tenure and relative tenure within the firm: the most junior (and least costly) workers are cut first, consistent with last-in, first-out patterns documented by Buhai et al. (2014). After layoffs, the firm has less incentive to cut the wages of surviving juniors than of seniors, whose average quality has not improved. The model can therefore replicate the tenure-gradient in layoffs and wage rigidity documented in the data.

The no–within-cohort discrimination assumption is central. The quoted intuition—"Layoffs give better control over who leaves the firm"—requires that layoffs be more useful than wage cuts for removing undesirable matches. Under complete information about match quality, layoffs would play no special role and the model would not generate meaningful layoffs. I offer several justifications for the assumption. One interpretation is nondiscrimination law: if firms cannot justify why they prefer one worker to another, they cannot selectively change wages. This is distinct from observable differences in worker quality. There is no doubt about productivity differences when Ann publishes significantly more than Bob, and the model need not restrict contracts based on observable quality. It is enough to assume the existence of some firm preferences over workers that are not known to workers themselves.

In the asymmetric-information case, I also consider an extension in which the firm may wage-discriminate. There remains a reason not to: as long as workers do not know their own quality, they effectively "risk-share" the layoff risk that arises. This makes layoffs cheaper for the firm, yielding two opposing instruments for removing poor matches. In Appendix A.3, I analyze a signaling game in which wages affect workers' beliefs about their type. Under a sufficiently inelastic probability of workers being retained (given outside opportunities), a pooling equilibrium exists in which the firm does not reveal match types. Focusing on this equilibrium justifies the maintained assumption.

2 Motivating evidence

I begin by documenting the relationship between layoffs and wage cuts using French matched employer–employee data. I examine how layoff rates and the wage pass-through of firm-level productivity shocks vary across firms and across worker tenure, and find substantial heterogeneity along both dimensions, irreconcilable with theories of exogenous wage rigidity. I use these facts as testable implications of my model.

2.1 Data

I use administrative data from France between 2009 and 2019. The four key variables for the analysis—wages, layoffs, productivity, and tenure—are either directly available or can be constructed given the richness of the data. I combine a worker panel from social-security records covering one-twelfth of the French labor force (providing wages, layoffs, and tenure) with annual firm balance-sheet data (providing productivity). For the sample, I focus on prime-age workers (25–55) in private-sector jobs with wages at least 5% above the national minimum wage at the time. Appendix C.1 provides further details on sample selection. The remaining sample contains 265,000 unique firms and 880,000 unique workers per year. I next describe how I construct the four key variables.

I measure labor productivity using value added per worker, as reported in the balance-sheet data. I model labor productivity y_{fst} at firm f in sector s at time t as

$$\log y_{fst} = \log a_t + \log b_{st} + \log x_{fst},$$

where a_t is the aggregate component, b_{st} is a sectoral component, and x_{fst} is a firm-level component. I residualize $\log y_{fst}$ on time dummies to extract the common time component, measure the sectoral component $\log b_{st}$ as the average productivity within a sector, and compute the firm component $\log x_{fst}$ as the residual. In what follows, I focus on firms' responses to the firm-specific component x_{fst} to abstract from broader general-equilibrium effects.

I measure wages as annual, CPI-adjusted labor earnings divided by days worked. Labor earnings are net of payroll taxes but pre–income tax, and include all forms of compensation (including bonuses and payments in kind) but exclude stock options. I residualize log wages on occupation, firm, and region dummies, as well as a quadratic in worker experience. I focus on job-stayer wage growth $\Delta \log w_{ift}$ for workers continuously employed at firm f in years $t-1$ and t. After computing growth rates for both productivity and wages, I trim the bottom and top 5% of each year's distributions.

I measure layoffs as breaks in employment spells of at least four weeks. The idea is that job-to-job transitions rarely entail long breaks, and given the low job-finding rate in France, recently laid-off workers are unlikely to find employment within a month. There remains a risk of both false positives and false negatives, as well as the risk of misclassifying voluntary transitions into medium-term nonemployment as layoffs. In Appendix C.2 I use the French Labor Force Survey to measure layoffs as quarterly movements from employment into unemployment, as self-reported by workers.

Lastly, workers' tenure at the firm is directly observed in the data. I focus on the first five years of tenure; beyond that, differences across cohorts are small.

	Layoff rate	Wage change
Low layoff rate	0.05%	0.017^{***}
		(0.0006)
Medium layoff rate	1.5%	0.012^{**}
		(0.0008)
High layoff rate	9.7%	-0.002^{***}
		(0.0006)

Table 1: Wage pass-through across firms. Data: DADS Panel + FARE, 2009–2019.

2.2 Wages and layoffs across firms

I begin by confirming the existing finding that firms exhibiting the most rigid wages lay off the most workers (see Ehrlich and Montes (2024)). To facilitate comparison, I group firms into terciles $d \in D$ based on their average layoff rates. For each group, I estimate the response of wage growth to firm productivity shocks. Define the growth rate of residualized wages for worker i in firm f between years t and $t-1$ as $\Delta \log w_{ift}$ and the growth rate of firm productivity as $\Delta \log(x_{ft})$. I estimate:

$$\Delta \log w_{ift} = \sum_{d \in D} \mathbf{1}\{f \in d\}\big(\alpha^d + \beta^d \Delta \log(x_{ft})\big) + \epsilon_{ift}.$$

Average layoff rates and estimated wage pass-throughs across firms are reported in Table 1. Firms in the low-layoff tercile raise wages by 1.7% in response to a 100% productivity shock. Firms in the middle tercile, with an average layoff rate of 1.5%, raise wages by 1.2%. Lastly, high-layoff firms, averaging 9.7% layoffs, reduce wages by 0.2% in response to a positive shock.

These results show that firms laying off the largest share of workers also exhibit the least responsive—and even negatively responsive—wages. This pattern is consistent with both the classical account (firms resort to layoffs when they cannot cut wages) and the mechanism developed here.

Wages and layoffs across tenure

Beyond cross-firm heterogeneity, the connection between rigid wages and layoffs also appears within firms, across worker tenure. Prior work has documented heterogeneity in layoff rates (e.g., Buhai et al. (2014)). Here I also document heterogeneity in wage pass-through.

Let workers' tenure be $ten \in T$, observed directly in the employer–employee data. Unlike the firm-level analysis, I consider each worker cohort up to five years of tenure. I run two regressions. First, analogous to the across-firm case, I estimate the response of wages to firm-level shocks across tenure:

$$\Delta \log w_{ift} = \sum_{ten \in T} \mathbf{1}\{ift \in ten\}\big(\alpha^{ten} + \beta^{ten}\, \Delta \log(x_{ft})\big) + \epsilon_{ift}. \tag{1}$$

Second, I estimate layoff rates using the individual layoff event EU_{ift} for worker i in firm f at year t:

$$EU_{ift} = \sum_{ten \in T} \mathbf{1}\{ift \in ten\}\, \alpha^{ten} + \epsilon_{ift}.$$

I also examine asymmetry in pass-through. Negative productivity shocks are when firms are most inclined to use wage cuts and layoffs. Although downward wage rigidity is well established (e.g., Hazell and Taska (2021)), its heterogeneity across tenure is less explored. To estimate it, I modify (1) by interacting productivity growth with an indicator for negative shocks, letting $\Delta \equiv \Delta \log(x_{ft})$:

$$\Delta \log w_{ift} = \sum_{ten \in T} \mathbf{1}\{ift \in ten\}\Big(\alpha^{ten} + \beta^{ten}\, \Delta + \tilde{\beta}^{ten}\, \mathbf{1}\{\Delta < 0\}\, \Delta\Big) + \epsilon_{ift}.$$

Table 2 reports the results. Junior workers exhibit the smallest average pass-through. At the same time, senior workers face layoff rates up to 5.5 times lower than juniors. Regarding downward rigidity, juniors benefit most: their response to positive shocks is largest, whereas their response to negative shocks is near zero. Later cohorts display much less asymmetry in pass-through and are therefore more exposed to pay cuts than juniors.

I interpret these findings as evidence of a cohort-level trade-off between wage cuts and layoffs: whenever and wherever firms lay off workers, they cut survivors' wages less. The heterogeneous treatment of cohorts cannot be explained by standard stories of wage rigidity (minimum wages, sectoral bargaining, morale costs) without additional assumptions. Likewise, severance payments in France rise only slightly with tenure (an extra payment worth $\approx$ 20 hours of pay per year of seniority).

2.3 Senior/junior wage gap response to layoffs

I shift the angle of analysis and examine how the wage gap between senior and junior workers responds to layoffs. Unlike earlier analyses, I neither study the overall cross-section nor condition on productivity shocks. The goal is to examine how layoffs—typically concentrated among juniors—affect within-firm wage inequality.

For each firm–year, I compute the median tenure and take the ratio of wages for workers above and below the median, $\bar{w}_{sen,f,t}/\bar{w}_{jun,f,t}$. I then compute the log change in this ratio

	Layoff rate	Avg. wage pass-through	Response to pos. shock	Response to neg. shock
<1 year	11%***	0.000	0.022***	0.002***
	(0.0002)	(0.004)	(0.004)	(0.003)
1–2 years	5%***	0.004	0.014***	0.020***
	(0.0002)	(0.004)	(0.004)	(0.003)
2–3 years	3%***	0.008**	0.013**	0.020***
	(0.0002)	(0.002)	(0.002)	(0.003)
3–4 years	2%***	0.008*	0.012*	0.015***
	(0.0002)	(0.003)	(0.003)	(0.004)
4–5 years	2%***	0.017***	0.010***	0.024***
	(0.0002)	(0.004)	(0.004)	(0.004)

Table 2: Layoffs and wage pass-through across tenure. Columns 3 and 4 report asymmetric pass-through to positive and negative shocks. Data: DADS Panel + FARE, 2009–2019.

	Layoff
Intercept	0.0031***
	(0.0001)
EU	−0.0036***
	(0.0006)

Table 3: Senior–junior wage log change in response to layoffs. Data: DADS Panel + FARE, 2009–2019.

over time and regress it on layoffs:

$$\Delta \log\bigl(\bar{w}_{sen,f,t}/\bar{w}_{jun,f,t}\bigr) = EU_{ft} + \epsilon_{ft}.$$

Estimates appear in Table 3. On average, the wage ratio rises over time (likely reflecting new hires), but it falls when layoffs occur, suggesting that surviving juniors experience faster wage growth than surviving seniors. These results are consistent with the interpretation that firms lay off junior workers and then cut the wages of surviving juniors by less than those of surviving seniors.

3 Model

I present a model of a frictional labor market in which firms sign workers of varying match quality to *dynamic* contracts. The model is designed to capture the heterogeneity in wage pass-through and layoffs across firms and tenure documented in the data.

3.1 Environment

Time is discrete and indexed by t. The economy is populated by a continuum of firms of measure 1, indexed by $j \in [0,1]$, and a continuum of workers of measure I, indexed by $i \in [0, I]$. Both types of agents are ex ante homogeneous and infinitely lived, with time-separable preferences and discount factor β. Firms are owned by outside investors who diversify firm-specific productivity risk. Thus firms maximize

$$E_0 \sum_{t=0}^{\infty} \beta^t \pi_{jt}.$$

Workers are risk-averse and have no access to financial markets. They consume home production b when unemployed and wage w when employed. Their utility is

$$E_0 \sum_{t=0}^{\infty} \beta^t u(c_{it}), \qquad u(c) = \frac{c^{1-\sigma}}{1-\sigma}.$$

Production

Firms may pay κ_e to start producing and must pay κ_f each period to remain open. Open firms employ a measure[1] n of workers. Each worker–firm match is either high or low quality and remains so for the duration of the match. Only the firm observes the quality of an individual match; the proportion of high-quality matches in the firm, z, is common knowledge. Production exhibits decreasing returns to scale in size n and, potentially, in quality z.[2] Production is also subject to firm-level shocks $y \in \mathcal{Y}$. Output depends on the numbers of high- and low-quality matches, n_H and n_L, equivalently on n and z:

$$yF(n_H, n_L) \equiv yF(n, z), \qquad n = n_H + n_L, \quad z = \frac{n_H}{n_H + n_L}.$$

Labor market

Each period, entering firms immediately hire $\tilde{n} = 1$ workers; incumbents hire $\tilde{n} \geq 0$. Workers—both employed and unemployed—search for jobs. Matching occurs in a frictional labor market with directed search, as in Moen (1997). There is a continuum of submarkets indexed by the promised value v to the worker. Firms choose in which submarket to post vacancies at cost c, and workers choose where to search. Within each submarket, matches

[1] The law of large numbers applies and is used throughout (Sun and Zhang (2009)).

[2] In the quantitative analysis, I restrict attention to decreasing returns in quality-adjusted quantity $g(z)n$.

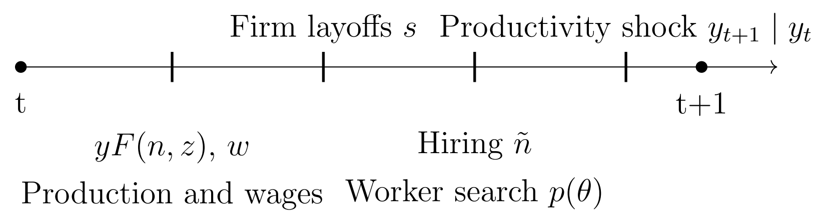


Figure 1: Within-period timeline

are formed according to a constant-returns-to-scale matching function; market tightness θ_v suffices to determine matching probabilities. Let $p(\theta_v)$ and $q(\theta_v) \leq 1$ denote the job-finding and vacancy-filling probabilities.

Firms are not restricted to a discrete number of vacancies and can deterministically hire $\tilde{n}$ workers from submarket v at cost $\tilde{n}\, c/q(\theta_v)$. The probability that a newly hired match is high quality is $z_0 \in [0,1]$, constant across agents and time. Upon hiring, the firm commits to deliver expected discounted utility v. The hiring trade-off is between the cost of hiring, $c/q(\theta_v)$, and the (higher) cost of employing a worker as v rises. Firms can downsize by laying off a fraction $s \in [0,1]$ of their workforce and by incentivizing incumbents to find jobs elsewhere.

I consider two versions of the economy: a steady state and, under an additional assumption (Appendix A.5), a block-recursive equilibrium following Menzio and Shi (2011) and Schaal (2017). In both cases, agents need not track the aggregate cross-sectional distribution.

Timing

Each period has four stages (Figure 1). First, production occurs: the firm collects output and pays wage w to each worker. Next, the firm lays off a fraction $s \geq 0$ of its workforce. Fired workers become unemployed and cannot search until next period. Then all workers—employed and unemployed—search, and all firms (entrants and incumbents) hire $\tilde{n}$. Hiring and search choices occur before next-period productivity y_{t+1} is realized, so agents take expectations $E_{y_{t+1}|y_t}$.

Information structure and contracts

Upon hiring, the firm commits to deliver expected utility v via a contract. A contract specifies wages and actions for the matched pair for all future firm-productivity histories $y^\tau \equiv (y_1, \ldots, y_\tau) \in \mathcal{Y}^\tau$. Firm productivity histories are common knowledge and therefore fully contractible. By contrast, match-specific productivity z_{ij} is private to the firm, and the worker's search decision $\hat{v}$ is private to the worker. The contract is

$$\mathcal{C} = \{w_\tau,\, s_\tau,\, \hat{v}_\tau\}_{\tau=t}^{\infty}. \tag{2}$$

Here w is the wage policy for each future productivity history. The second component, s, is the *expected* layoff probability from the worker's perspective (who does not observe their match quality). These probabilities are history-dependent: in histories where information about match quality is updated, future layoff probabilities reflect the worker's Bayesian update. For example, after multiple negative shocks that force large layoffs, remaining workers rationally assign higher probability to being high quality, and subsequent layoff probabilities adjust accordingly. The last component is the worker's (unobserved) search decision. I focus on contracts in which recommended search is incentive compatible, i.e., the contract specifies workers' search choices subject to the constraint that those choices are optimal for workers.

The contract space allows fully flexible wage and layoff responses to productivity histories. With a continuum of concurrent contracts, the firm can choose how to treat a heterogeneous workforce (by quality and contract): when a negative shock hits, whom to fire and whose wages to cut. This feature is central and specific to the setting: unlike models with CRS production, these decisions depend on the entire firm state; unlike dynamic models with Nash bargaining (McCrary (2022)) or sequential bargaining (Bilal et al. (2022)), workers within the same firm may optimally face different wages, layoff risks, and responses to shocks.

3.2 Value functions

The contract and all agents' problems admit a recursive formulation. I begin with workers' problems and then turn to firms managing a continuum of contracts. I show that the firm's problem can be reformulated with a discrete state space.

Worker's problem

Unemployed workers consume b and each period choose the submarket that offers the best trade-off between promised future utility and job-finding probability. In a stationary equilibrium (suppressing time subscripts), the value of unemployment U is

$$U = \max_{v}\ u(b) + \beta\big[(1 - p(\theta_v))U + p(\theta_v)v\big]. \tag{3}$$

Consider an employed worker owed value v. Suppose the firm pays wage w this period, lays off with probability s, and promises future value v' from next period onward. The worker's search problem is

$$v = \max_{\hat{v}} \; u(w) + \beta\Big[sU + (1-s)\big((1-p(\theta_{\hat{v}}))v' + p(\theta_{\hat{v}})\hat{v}\big)\Big]. \tag{4}$$

The optimal search target $\hat{v}$ depends only on the promised v'. By raising v', the firm induces search in higher-$\hat{v}$ submarkets, lowering the probability that the worker exits. Equivalently,

$$v = u(w) + \beta\big[sU + (1-s)R(v')\big],$$

where $R(v') \equiv \max_{\hat{v}} \big[(1-p(\theta_{\hat{v}}))v' + p(\theta_{\hat{v}})\hat{v}\big]$ is the worker's optimal continuation value given promise v' and no layoff.

Firm's problem

A firm employs a measure n of workers. Let $P(v)$ denote the distribution of promised values owed to incumbents. For each v, the firm chooses the wage w_v, layoff rate s_v, and next period's productivity-history–contingent promised values $\{v'_{v,y'}\}$. The firm may also hire $\tilde{n}$ workers at value $\tilde{v}$. Each match can be high or low quality. The firm cannot set quality-contingent wages; it affects match quality only through layoffs.[3] The recursive problem is

$$\begin{aligned}
J(y,n,P(v),z) = \max_{\tilde{n},\tilde{v},\{w_v,s_v,v'_{v,y'}\}} \; & yF(n,z) - \int w_v \, dP(v) - \tilde{n}\frac{c}{q(\tilde{v})} - \kappa_f + \beta\, E_{y'|y}\, J\big(y',n',P'(v)\big) \\
\text{s.t.} \quad & u(w_v) + \beta\big[s_v U + (1-s_v)R(v'_v)\big] = v \quad \forall v, \\
& v'_v = E_{y'|y}\, v'_{v,y'} \quad \forall v, \\
& n' = n\int (1-s_v)\big(1-p(v'_v)\big)\, dP(v) + \tilde{n}, \\
& n'P'(v) = n\int E_{y'|y}\, \mathbb{1}\{v'_{v,y'} \le v\}(1-s_v)\big(1-p(v'_v)\big)\, dP(v) + \mathbb{1}\{\tilde{v} \le v\}\tilde{n}.
\end{aligned}$$

The firm maximizes the present value of profits subject to honoring each worker's promised value. Because search occurs before the next productivity state is realized, workers care about the expected promised value v'_v rather than particular realizations $v'_{v,y'}$. The last two constraints describe the laws of motion for firm size and the distribution of promises.

Discretizing the problem As written, the state includes a probability distribution—an uncountably infinite-dimensional object. I show that the state space can be discretized, yielding a countably infinite state. First, under directed search, a firm posts in a single

[3] Appendix A.3 shows that, under sufficiently inelastic on-the-job search, this is a pooling equilibrium of a signaling game in which the firm *could* set quality-contingent wages.

submarket and hires at a single value $\tilde{v}$.[4] Hence, all workers hired in the same period by the same firm are owed the same expected utility—both at hiring and, given no quality-contingent wages, thereafter. It is therefore equivalent to work with the CDF $P(v)$ or the PMF $\mathbb{P}(V = v)$, with $P(v) = \sum_{v' \leq v} \mathbb{P}(V = v')$. For a firm of age $K < \infty$, there are at most K distinct promised values with positive probability, corresponding to cohorts by tenure $k = t - t_{\text{hired}} \leq K$. The state can thus be recast by tenure:

Lemma 1. *For a firm of age K, the problem $J(y, n, P(v))$ is equivalent to*

$$
\begin{aligned}
J\big(y, \{n_k, v_k, z_k\}_{k\leq K}\big) = \max_{\tilde{n}, \tilde{v}, \{v'_{y',k}, w_k, s_k\}_{k\leq K}} \quad & yF\Big(\sum_k n_k, \frac{\sum_k n_k z_k}{\sum_k n_k}\Big) - \sum_k w_k n_k - \tilde{n}\frac{c}{q(\tilde{v})} - \kappa_f \\
& + \beta\, E_{y'|y}\, J\big(y', \{n'_k, v'_k, z'_k\}_{k\leq K+1}\big) \\
s.t. \qquad & u(w_k) + \beta\big[s_k U + (1 - s_k) R(v'_{k+1})\big] = v_k \quad \forall k \leq K, \\
& v'_{k+1} = E_{y'|y}\, v'_{k+1,y'} \quad \forall k \leq K, \\
& n'_{k+1} = n_k(1 - s_k)\big(1 - p(v'_{k+1})\big) + \tilde{n} \quad \forall k \leq K, \\
& z'_{k+1} = \min\Big\{\frac{z_k}{1 - s_k}, 1\Big\} \quad \forall k \leq K, \\
& n'_0 = \tilde{n}, \quad v'_0 = \tilde{v}, \quad z'_0 = z_0.
\end{aligned}
$$

This representation both discretizes the state space and clarifies how layoffs and wages interact within cohorts. Although workers do not know their own match quality, they know the cohort-level share of high-quality matches. New hires start with z_0, and while z_k may evolve, all workers of the same tenure k share the same probability $z_k \geq z_0$ of being high quality. This probability, common knowledge to the firm and workers, depends on layoffs in the cohort. As shown below, average cohort quality has a first-order effect on wage growth, linking wage pass-through and layoffs within the firm.

Free entry and exit

Firms enter by paying κ_e. Upon entry, a firm draws productivity and starts with a single worker. Free entry pins down expected profits at entry:

$$
\kappa_e \geq \max_{v_0} \Big(- \frac{c}{q(\theta_{v_0})} + \beta\, E_y\, J\big(y, \{1, 0, \ldots\}, \{v_0, \ldots\}, \{z_0, \ldots\}\big)\Big). \tag{5}
$$

Empirically, this implies a negative relationship between entry costs κ_e and vacancy-filling probabilities $q(\theta)$: cheaper entry tightens the labor market, raising job-finding rates and lowering vacancy-filling probabilities.

[4] I rule out mixed strategies by an individual firm: if indifferent across submarkets, it posts in one of them.

Incumbents pay an operating cost κ_f each period to stay open. With κ_f already in J, firms remain open if

$$J\big(y, \{n_k, v_k, z_k\}\big) \geq 0. \tag{6}$$

3.3 Equilibrium

An equilibrium is a sequence of policies, matching rates, and distributions of workers and firms across submarkets v such that, each period:

- Firms solve the problem in Lemma 1.
- Workers solve (3) and (4).
- Free-entry and free-exit conditions (5)–(6) hold.
- Job-finding and vacancy-filling probabilities are consistent with the matching function.
- Tightness θ_v is consistent with firms' posting and workers' search strategies.
- The labor market clears.

Under the assumption in Appendix A.5, the equilibrium may be block recursive—independent of the aggregate distributions. I use that assumption in the quantitative analysis, but not in the theoretical discussion, where I focus on the steady state described above.

3.4 Mechanism

I show how the model generates the empirical connection between layoffs and wage pass-through, and the associated tenure heterogeneity.

Wage growth

Proposition 1. *For any current state $(y, \{n_k, v_k, z_k\})$, wages evolve according to*

$$\frac{1}{u'(w'_{k+1})} - \frac{1}{u'(w_k)} = \eta(v'_{k+1})\, E_{y'|y}\, \frac{\partial J\big(y', \{n'_k, v'_{y',k}, z'_k\}\big)}{\partial n'_{k+1}},$$

where $\eta(v'_{k+1}) \equiv \frac{\partial \log(1-p(v'))}{\partial v'}\big|_{v'=v'_{k+1}}$ is the semi-elasticity of the job-finding probability with respect to the promised value.

Proof. See Appendix A.1. $\square$

This relationship captures the insurance–incentives trade-off. When the marginal value of a worker, $E_{y'|y}\big[\partial J/\partial n'_{k+1}\big]$, is positive, the firm prefers to retain workers and backloads wages (raising v'). When the marginal value is negative, the firm lowers wages to encourage quits. This motivates a target wage:

Definition 1. *The target wage for cohort k, $w^*_k\big(y, \{n_k, v_k, z_k\}/v_k\big)$, is the wage associated with the promised value $v^*_k\big(y, \{n_k, v_k, z_k\}/v_k\big)$ that solves*

$$M\big(y, \{n_k\}, v_0, \ldots, v_k, \ldots, \{z_k\}\big) \equiv E_{y'|y} \frac{\partial J\big(y', \{n'_k, v'_k, z'_k\}\big)}{\partial n'_{k+1}} = 0,$$

where $\{n'_k, v'_k, z'_k\}$ are next-period states implied by Lemma 1. The target wage equals

$$w^*_k\big(y, \{n_k, v_k, z_k\}/v_k\big) = u'^{-1}\left(- \frac{n'_{k+1}}{\partial J\big(y', \{n'_k, v'_k, z'_k\}\big)/\partial v'_{y',k+1}} \right).$$

At the target, the cohort's marginal profit is zero, making the firm indifferent to poaching (while still preferring to shed low-quality matches within the cohort). The target wage governs within-firm wage movements:

Proposition 2. *For any state $(y, \{n_k, v_k, z_k\})$ and cohort k there exists a target wage w^*_k such that:*

1. *Wages move toward the target:*

$$w_k \le w^*_k \;\Rightarrow\; w_k \le w'_{k+1} \le w^*_k, \qquad w_k \ge w^*_k \;\Rightarrow\; w^*_k \le w'_{k+1} \le w_k.$$

2. *The farther from target, the faster the adjustment:*

$$|w_k - w^*_k| \ge |w_{k'} - w^*_{k'}| \;\Rightarrow\; |w'_{k+1} - w_k| \ge |w'_{k'+1} - w'_{k'}|.$$

   To first order:

3. *w^*_k increases in z_k and is unchanged in $z_{k'}$ for $k' \ne k$.*

4. *All target wages respond equally to other state variables:*

$$\frac{\partial w^*_k}{\partial s} = \frac{\partial w^*_{k'}}{\partial s} \quad \forall k, k',\; s \ne z_k, z_{k'}.$$

Proof. See Appendix A.1. □

Thus, at any state and for any cohort, wages adjust toward w^*_k. Layoffs are the firm's only instrument to alter within-cohort quality; by item (4), they are also the only way to differentially move targets across cohorts.

Proposition 3. *Consider a state* $(y, \{n_k, v_k\})$ *and a cohort* k *with* $z_k < 1$*. Optimal layoffs satisfy*

$$- E_{y'|y} \frac{\partial J'}{\partial n'_{k+1}} \big(1 - p(v'_{k+1})\big) + E_{y'|y} \frac{\partial J'}{\partial z'_{k+1}} \frac{\partial z'_{k+1}}{\partial s_k} \frac{1}{n_k} - \frac{R(v'_{k+1}) - U}{u'(w_k)} \;\leq\; 0,$$

with $s_k \geq 0$ *and complementary slackness. Moreover, cohorts with lower promised values are more exposed to layoffs:*

$$v_k \leq v_{k'} \;\Rightarrow\; s_k(1 - z_k)n_k \;\geq\; s_{k'}(1 - z_{k'})n_{k'} \quad \forall\, k' \leq K,\; z_{k'} < 1.$$

To first order, layoff decisions load equally on quality across cohorts:

$$\frac{\partial s_k}{\partial z_k} = \frac{\partial s_k}{\partial z_{k'}} \quad \forall\, k' \leq K.$$

Proof. See Appendix A.1. □

As with wages, layoffs trade off the marginal value of a worker, $E_{y'|y}[\partial J'/\partial n'_{k+1}]$, against the compensation cost $\big(R(v'_{k+1}) - U\big)/u'(w_k)$. The distinctive term is the quality effect, $E_{y'|y}[\partial J'/\partial z'_{k+1}]\,(\partial z'_{k+1}/\partial s_k)/n_k$, since quality cannot be priced but can be *selected* via layoffs. Lower-v cohorts are more exposed to layoffs regardless of quality (provided $z < 1$). Barring large quality gaps, these lower-v cohorts also tend to be farther below their targets and, after layoffs, exhibit faster wage growth.

Finally, to map these results to tenure patterns, note that higher-tenure workers generally have higher promised values than juniors due to wage backloading (Proposition 2, item 1). Juniors, being owed less, are therefore more exposed to layoffs than seniors and, via the upward shift in their target wages induced by layoffs, experience smaller wage cuts (if any) than seniors.

4 Quantitative Analysis

I calibrate the model using administrative data from France and assess whether it reproduces the empirical evidence on wage pass-through and layoffs across firms and tenure.

4.1 Solving the model

The tenure-based formulation yields a discrete—though expanding—state space. To make the problem fully tractable, I note that wages in contracts tend to converge, a result that follows from the propositions above. A corollary of Proposition 2 is that wages for cohorts with the same quality converge because they share the same target wage. I also expect quality across cohorts to converge:

- Lower-quality cohorts (relative to others) have lower target wages;
- Over time, low-quality cohorts become low-value cohorts;
- By Proposition 3, low-value cohorts are laid off first;
- Hence lower-quality cohorts catch up in quality and, consequently, in wages.

Given convergence, the practical question is how quickly it occurs—or, equivalently, how many cohorts to track. Empirically (Appendix C.3), wage growth in France flattens after about ten years of tenure. I therefore restrict attention to a finite and constant $K \leq 10$ for all firms in the quantitative analysis. This is an approximation: as $K \to \infty$, the model converges to the problem in Lemma 1.

A second complication, common in dynamic-contract models, is the large action space because $\{v'_{y',k}\}$ grows with the number of productivity states. Appendix A.4 shows that the problem can be solved in its dual form by choosing future marginal utilities—constant across productivity realizations—instead of promised values.

Lastly, I solve for a block-recursive equilibrium. As shown in Appendix A.5, such an equilibrium exists when all new hires enter at the same promised value v_0; any additional value associated with higher submarkets v is paid as a sign-on wage. I set v_0 equal to the unemployment value U.

4.2 Model specification

I work at annual frequency: contracts in France are rarely updated more than once per year, which facilitates cohort tracking. I set the discount factor to $\beta = 0.96$, consistent with a 4% annual interest rate.

I use logarithmic utility so that the composition $u' \circ u^{-1}(v)$ is strictly increasing; within CRRA preferences, this property holds only for log utility.

Production is a concave function of quality-adjusted quantity:

$$F(n, z) = [n\,(z + \alpha_z(1 - z))]^{\alpha}.$$

I set $\alpha = 0.85$, consistent with recent work on firm dynamics (Schaal (2017); Bilal et al. (2022)).

Matching follows a CES contact-rate function (as in Menzio and Shi (2011)):

$$p(\theta) = \theta\,(1 + \theta^{\gamma})^{-1/\gamma}, \qquad q(\theta) = (1 + \theta^{\gamma})^{-1/\gamma}.$$

Idiosyncratic firm-level productivity follows a uniform Markov process as in Balke and Lamadon (2022): with probability λ_y productivity is redrawn uniformly; otherwise it remains unchanged.

The remaining parameters to estimate are:

- Unemployment production b;
- Productivity: persistence λ_y and variance σ_y;
- Search: vacancy-posting cost c, matching curvature γ, and on-the-job (OJS) search efficiency λ_{jj};
- Firm dynamics: entry cost κ_e and operating cost κ_f;
- Match heterogeneity: share of high-quality matches at hire q_0 and relative productivity of low-quality matches α_z.

4.3 Moments of interest

I target four sets of moments: transition probabilities, wage growth, productivity dynamics, and firm dynamics.

Transitions. I measure annual job-to-job (E2E) and unemployment-to-employment (U2E) transitions. For the hiring rate, I use the share of newly hired workers among all observations. U2E and E2E map directly to c and λ_{jj}, respectively. I obtain an E2E rate of 6.3% and a 12.8% share of new hires economy-wide.

Layoffs across productivity. I track the distribution of employment-to-unemployment (E2U) rates across firms by productivity. These moments discipline match-heterogeneity parameters: the lower the relative productivity of low-quality matches α_z, the more even highly productive firms will lay off workers. After sufficiently positive productivity shocks, firms hire; when q_0 (the share of high-quality new matches) is small, subsequent layoffs are more likely. I find little difference in layoff rates across productivity groups, with the range of 1 percentage point between top and the bottom tercile.

Wage growth by tenure. I use wage growth over the first five years of tenure to discipline b, following Souchier (2022). Intuitively, wages rise until a cohort's marginal profit reaches zero; the lower the cohort's starting point (which depends on the unemployment outside option), the larger the initial wage growth. I document 6.1% wage growth after five years of tenure.

Productivity dynamics. I measure the standard deviation and persistence of firm-level productivity (net of aggregate and sectoral effects) to discipline λ_y and σ_y. Firm-level productivity is moderately persistent with coefficient 0.79. Its standard deviation is 0.39, the largest among productivity components, consistent with comparable estimates of 0.81 and 0.30 in Souchier (2022).

Firm dynamics. To discipline κ_e, I target a 6% share of jobs created by entering firms. Because κ_f affects exit and thus selection, I target an average firm size of 32.5. For comparison, the average establishment size is 17.9, close to 15.6 in the 2002 U.S. Economic Census.

4.4 Untargeted moments

As validation, I assess whether the model reproduces the heterogeneity in layoffs and wage pass-through across firms and tenure documented in Section 2.

Work in progress.

5 Conclusion

References

1. Balke, Neele and Thibaut Lamadon (2022). "Productivity shocks, long-term contracts, and earnings dynamics". In: *American Economic Review* 112.7, pp. 2139–2177.

2. Bennedsen, Morten et al. (Feb. 2025). *Why Firms Lay Off Workers Instead of Cutting Wages: Evidence from Linked Survey-Administrative Data*. Working Paper Series 2025-05. Federal Reserve Bank of San Francisco. DOI: 10.24148/wp2025-05. URL: https://ideas.repec.org/p/fip/fedfwp/99558.html.

3. Bilal, Adrien et al. (2022). "Firm and worker dynamics in a frictional labor market". In: *Econometrica* 90.4, pp. 1425–1462.

4. Buhai, I Sebastian et al. (2014). "Returns to tenure or seniority?" In: *Econometrica* 82.2, pp. 705–730.

5. Cho, In-Koo and David M. Kreps (May 1987). "Signaling Games and Stable Equilibria*". In: *The Quarterly Journal of Economics* 102.2, pp. 179–221. ISSN: 0033-5533. DOI: 10.2307/1885060. eprint: https://academic.oup.com/qje/article-pdf/102/2/179/5441738/102-2-179.pdf. URL: https://doi.org/10.2307/1885060.

6. Davis, Steven J. and Pawel M. Krolikowski (Feb. 2025). "Sticky Wages on the Layoff Margin". In: *American Economic Review* 115.2, pp. 491–524. DOI: 10.1257/aer.20240309. URL: https://www.aeaweb.org/articles?id=10.1257/aer.20240309.

7. Ehrlich, Gabriel and Joshua Montes (Jan. 2024). "Wage Rigidity and Employment Outcomes: Evidence from Administrative Data". In: *American Economic Journal: Macroeconomics* 16.1, pp. 147–206. DOI: 10.1257/mac.20200125. URL: https://www.aeaweb.org/articles?id=10.1257/mac.20200125.

8. Hazell, Jonathon and Bledi Taska (2021). *Downward rigidity in the wage for new hires*. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3728939.

9. McCrary, Sean (2022). "A Job Ladder Model of Firm, Worker, and Earnings Dynamics". In: *Worker, and Earnings Dynamics (November 4, 2022)*.

10. Menzio, Guido and Shouyong Shi (2011). "Efficient search on the job and the business cycle". In: *Journal of Political Economy* 119.3, pp. 468–510.

11. Moen, Espen R (1997). "Competitive search equilibrium". In: *Journal of political Economy* 105.2, pp. 385–411.

12. Schaal, Edouard (2017). "Uncertainty and unemployment". In: *Econometrica* 85.6, pp. 1675–1721.

13. Souchier, Martin (2022). *The pass-through of productivity shocks to wages and the cyclical competition for workers*. Tech. rep. Working Paper.

14. Sun, Yeneng and Yongchao Zhang (2009). "Individual risk and Lebesgue extension without aggregate uncertainty". In: *Journal of Economic Theory* 144.1, pp. 432–443. ISSN: 0022-0531. DOI: `https://doi.org/10.1016/j.jet.2008.05.001`. URL: `https://www.sciencedirect.com/science/article/pii/S0022053108000768`.

A Model Appendix

A.1 Proofs

Proof of Proposition 1

The proposition is derived from the first-order conditions of the firm's problem.

Proof. Consider the problem from Lemma 1. The FOC for v'_{k_1} yields (denote ρ_k the shadow cost of PK_k and ω_k the shadow cost of the expectation condition):

$$\rho_k \beta(1-s_k)(1-p(v'_{k+1})) - \omega_k + \beta n_k(1-s_k)\frac{\partial(1-v(v'_{k+1}))}{\partial v'_{k+1}} E_{y'|y} \frac{\partial J(y', \{n'_k, v'_{y',k}\}_{k\leq K+1})}{\partial n'_{k+1}} = 0$$

To develop this further, I use the FOC for w_k:

$$-n_k + \rho_k u'(w_k) = 0 \iff \rho_k = \frac{n_k}{u'(w_k)}$$

One can similarly develop ω_k by applying the FOC for $v'_{y',k+1}$:

$$\beta Pr_{y'|y} \frac{\partial J(y', \{n'_k, v'_{y',k}\}_{k\leq K+1})}{\partial v'_{y',k+1}} + Pr_{y'|y}\omega_k = 0 \iff \omega_k = -\beta \frac{\partial J(y', \{n'_k, v'_{y',k}\}_{k\leq K+1})}{\partial v'_{y',k+1}} \forall y'$$

Applying the envelope theorem one can then note that

$$-\frac{\partial J(y', \{n'_k, v'_{y',k}\}_{k\leq K+1})}{\partial v'_{y',k+1}} = \frac{n'_{k+1}}{u'(w'_{k+1})}$$

Putting all this together, we get

$$\frac{n_k}{u'(w_k)}\beta(1-s_k)(1-p(v'_{k+1})) - \beta\frac{n'_{k+1}}{u'(w'_{k+1})} + \beta n_k(1-s_k)\frac{\partial(1-v(v'_{k+1}))}{\partial v'_{k+1}} E_{y'|y} \frac{\partial J(y', \{n'_k, v'_{y',k}\}_{k\leq K+1})}{\partial n'_{k+1}} = 0$$

All that is left is to note that $n'_{k+1} = n_k(1-s_k)(1-p(v'_{k+1}))$, divide by the common terms $(\beta n'_{k+1})$ and rearrange.

$$\frac{1}{u'(w'_{k+1})} - \frac{1}{u'(w_k)} = \eta(v'_{k+1}) E_{y'|y} \frac{\partial J(y', \{n'_k, v'_{y',k}\})}{\partial n'_{k+1}}$$

□

Proof of Proposition 2

Proof. □

Proof of Proposition 3

I start by describing the FOC with respect to layoffs s_k:

Proof. Consider the case where $s_k > 0$. Then

$$-\beta n_k(1-p(v'_{k+1}))E_{y'|y}\frac{\partial J(y', \{n'_k, v'_{y',k}, z'_k\})}{\partial n'_{k+1}} + \beta E_{y'|y}\frac{\partial J(y', \{n'_k, v'_{y',k}, z'_k\})}{\partial z'_{k+1}}\frac{\partial z'_{k+1}}{\partial s_k} - \rho_k\beta(U - R(v'_{k+1})) = 0$$

Note that the FOC with respect to w_k yields $\rho_k = \frac{n_k}{u'(w_k)}$ and divide by βn_k to get the FOC in the Propositon.

For the further results, I rewrite the firm state $(y, \{n_k, v_k, z_k\})$ into $(y, \{\underline{n}_k, \bar{n}_k, v_k\})$, where $\bar{n}_k = z_k n_k$ and $\underline{n}_k = (1-z_k)n_k$.

I focus on the case where $\underline{n}_k > 0$ and $s_k(\underline{n}_k + \bar{n}_k) \leq \underline{n}_k$, that is, firm has not yet fired all the bad matches with the cohort (in the previous FOC, this is equivalent to $\frac{\partial z'_{k+1}}{\partial s_k} > 0$). The marginal value of firing a worker is then

$$-\frac{R(v'_{k+1}) - U}{u'(w_k)} - (1-p(v'_{k+1}))E_{y'|y}\frac{\partial J(y', \{\underline{n'}_k, \bar{n'}_k, v'_k\})}{\partial \underline{n}'_{k+1}}$$

To show that cohorts with lower promised values are more subject to layoffs, I take a derivative of this FOC with respect to the promised value v_k.The only component in the FOC directly dependent on v_k is $\frac{1}{u'(w_k)}$, where $w_k = u^{-1}(v_k - \beta[s_k U + (1-s_k)R(v'_{k+1})])$. Due to the $CRRA$ utility function, we find that $\frac{1}{u'(w_k)}$ is increasing in v_k, and, therefore, the marginal profit of firing workers is decreasing in v_k.

Lastly, I show that layoffs are equally dependent on the quality of any cohort (as long as $z_k < 1$). First, I note that quality directly appears only in the marginal value of low quality workers

$$\frac{\partial J(y', \{\underline{n}'_k, \bar{n}'_k, v'_k\})}{\partial \underline{n}'_{k+1}} = y'[F'_1(\sum n'_k, \frac{\sum n'_k z'_k}{\sum n'_k}) - F'_2(\sum n'_k, \frac{\sum n'_k z'_k}{\sum n'_k})\frac{\sum n'_k}{(\sum n'_k z'_k)^2} - w'_{k+1} + \beta E_{y''|y'}\frac{\partial J(y'', ...)}{\partial \underline{n}''_{k+2}}\frac{\partial \underline{n}''_{k+}}{\partial \underline{n}'_{k+}}$$

Next, I note that all the quality states $\{z_k\}$ contribute in the same manner to $\frac{\partial J(y',\{\underline{n}'_k,\bar{n}'_k,v'_k\})}{\partial \underline{n}'_{k+1}}$, no matter the cohort k. Therefore, to first-order, the marginal value of firing a worker is equally dependent on all the quality states.

I don't think I need to assume $s_k(\underline{n}_k + \bar{n}_k) \leq \underline{n}_k$. I can write the proof for both cases, and the results should be there for either one. □

Proof of Proposition ??

Start with the wage growth equation from Proposition 1, extended to the case of heterogeneous matches.

$$\frac{1}{u'(w'_{k+1})} - \frac{1}{u'(w_k)} = \eta(v'_{k+1})E_{y'|y}\frac{\partial J(y',\{n'_k, v'_{y',k}, z'_k\})}{\partial n'_{k+1}}$$

One can now apply the Envelope Theorem to extend the RHS of the equation:

$$E_{y'|y}\frac{\partial J(y',\{n'_k, v'_{y',k}, z'_k\})}{\partial n'_{k+1}} = E_{y'|y}\Big[y'\frac{\partial F(\sum n'_k, \frac{\sum n'_k z'_k}{\sum n'_k})}{\partial n'_{k+1}} - w'_{k+1} + \beta(1-p(v''_{k+2}))(1-s'_{k+1})E_{y''|y'}\frac{\partial J(y'',\{n''_k, v''_{y'',k}, z''_k\})}{\partial n''_{k+2}})\Big]$$

One can note that the production derivative $\frac{\partial F(\sum n'_k, \frac{\sum n'_k z'_k}{\sum n'_k})}{\partial n'_{k+1}}$ is increasing in z'_{k+1}. And, moreover, that $\frac{\partial^2 F(\sum n'_k, \frac{\sum n'_k z'_k}{\sum n'_k})}{\partial n'_{k+1}\partial z'_{k+1}} > \frac{\partial^2 F(\sum n'_k, \frac{\sum n'_k z'_k}{\sum n'_k})}{\partial n'_{k'}\partial z'_{k+1}}$, $k' \neq k+1$. This implies that, the workers in steps $k \in K_s$ will experience the largest rise in marginal productivity. This doesn't mean that they are at the highest marginal productivity though! Part of my intuition is predicated on the fact that the other workers, ones not fired, are already quite close to their target wages! Otherwise those guys would still have the best wage growth. Essentially, my intuition is that the fired cohorts receive a spike in marginal productivity, so, if they're on the similar-ish wage path to the other cohorts, they should get the best wage growth. But this needs to be formalized some more.

A.2 Tenure-specific Severance Payments

I allow the firm to offer tenure-specific severance payments sev_k to its workers. The severance is constant over time and paid perpetually upon firing and before finding a new job. I show that the severance structure involves higher payments for longer tenured workers (if those workers are on a higher promised value).

Proposition 4. *Fix a firm state. Its severance payments for each tenure k are given by*

$$\frac{u'(b+sev_k)}{u'(w_k)} = 1 - \frac{\beta sev_k \frac{\partial p(\theta_{sev_k})}{\partial sev_k}}{1-\beta(1-p(\theta_{sev_k}))}$$
$$\theta_{sev_k} = \theta(arg\max_v[(1-p(v))U(sev_k)+p(v)v])$$

Proof. I start by describing the unemployment value of a worker with severance payment sev_k:

$$U(sev_k) = u(b+sev_k) + \beta \max_v[(1-p(\theta_v))U(sev_k)+p(\theta_v)v]$$

Denote the probability of finding a job with severance payment sev_k as $p(\theta_{sev_k})$. The extra value to the unemployed from the severance payment is then given by

$$\frac{\partial U(sev_k)}{\partial sev_k} = u'(b+sev_k) + \beta(1-p(\theta_{sev_k}))U'(sev_k) = \frac{u'(b+sev_k)}{1-\beta(1-p(\theta_{sev_k}))}$$

Then the total benefit to the firm from raising the severance payment is the slackening of the promised-keeping constraint thanks to this rise in the unemployment value:

$$\lambda_k n_k \beta s_k \frac{\partial U(sev_k)}{\partial sev_k} = \frac{n_k}{u'(w_k)}\beta s_k \frac{u'(b+sev_k)}{1-\beta(1-p(\theta_{sev_k}))}$$

On the cost side, the firm internalizes the net present value of the severance payments when firing $n_k s_k$ workers:

$$\frac{\partial}{\partial sev_k}\left[n_k s_k \beta \frac{sev_k}{1-\beta(1-p(\theta_{sev_k}))}\right] = n_k s_k \beta \frac{[1-\beta(1-p(\theta_{sev_k}))] - \beta sev_k \frac{\partial p(\theta_{sev_k})}{\partial sev_k}}{[1-\beta(1-p(\theta_{sev_k}))]^2}$$

The optimal severance payment then follows from the first-order condition:

$$\frac{n_k}{u'(w_k)}\beta s_k \frac{u'(b+sev_k)}{1-\beta(1-p(\theta_{sev_k}))} = n_k s_k \beta \frac{[1-\beta(1-p(\theta_{sev_k}))] - \beta sev_k \frac{\partial p(\theta_{sev_k})}{\partial sev_k}}{[1-\beta(1-p(\theta_{sev_k}))]^2}$$

Rearranging gives the result.

$$\frac{u'(b+sev_k)}{u'(w_k)} = 1 - \frac{\beta sev_k \frac{\partial p(\theta_{sev_k})}{\partial sev_k}}{1-\beta(1-p(\theta_{sev_k}))}$$

□

Note that, besides $u'(w_k)$, all the components of the severance payment are independent of both the firm state and the worker tenure. It is immediate to notice then that higher paid workers will have higher severance payments: as $\frac{1}{u'(w_k)}$, the value to the firm of the severance payment goes up, while costs stay the same. Therefore, the firm will optimally choose to offer higher severance payments to higher paid workers.

This equation is also easy to implement numerically: using the formulation in Appendix A.4, where $\rho_k \equiv u'(w_k)$ is a state variable, I can immediately compute the payments for all the firm states, before solving the rest of the firm problem.

A.3 Microfounding the Wage Noncontractability

A version of my model where the firm is allowed to choose quality-specific wage is in fact a signalling game: firm's action of choosing the wage signals the workers their quality. In this section I show that a pooling Perfect Bayesian Equilibrium of such a game exists, and thus it is plausible that no information is conveyed.

As a full infinite-horizon, multiple-receivers, complex sender model is too complicated of a game to solve, I restrict attention to a simplified model.

Consider a 3-period version of the model. The firm starts with measure $n = 2$ of workers, half of them of high quality and the other half of low. I allow the firm to offer quality-contingent wages in whichever way it likes.

I show that, under a sufficiently small elasticity of job search probability with respect to promised value v', $\eta(v') \equiv \frac{\partial(1-p(v'))/\partial v'}{(1-p(v'))}$, the pooling Perfect Bayesian Equilibrium exists. Moreover, under a stronger condition on the elasticity, this PBE survives the Intuitive Criterion (Cho and Kreps (1987)).

A.4 Recursive Lagrangian Approach

The original design of the problem would require solving promised values $v'_{y',k}$ for both each tenure step and each future productivity state. Following Balke and Lamadon (2022), I solve the following Pareto problem:

$$\begin{aligned}
\mathcal{P}(y, \{n_k, \rho_k, z_k\}) = &\inf_{\omega_k} \sup_{\tilde{n}, \tilde{v}, \{w_k, s_k, v'_k\}} yF(n, z) - \sum_k n_k w_k - \kappa_f - \tilde{n}\frac{c}{q(\theta_{\tilde{v}})} \\
&+ \sum_k \rho_k (u(w_k) + \beta[s_k U + (1 - s_k) R(v'_{k+1})] \\
&- \beta \sum_k \omega_k v'_{k+1} + \beta E_{y'|y} \mathcal{P}(y', \{n'_k, \omega_k, z'_k\})
\end{aligned}$$

where

$$\mathcal{P}(y, \{n_k, \rho_k, z_k\}) \equiv \sup_{\{v_k\}} J(y, \{n_k, v_k, z_k\}) + \sum_k \rho_k v_k$$

The following proof (for $K \to \infty$ but the proof extends trivially to finite K) establishes its equivalence with the initial problem. It follows the steps of Balke and Lamadon (2022), extending it to the case of a multi-worker firm.

Proof. We have the following recursive formulation for J:

$$
\begin{aligned}
J(y,\{n_k,v_k,z_k\}_{k\leq K}) = \max_{\tilde{n},\tilde{v},\{v'_k,v'_{y',k},w_k,s_k\}_{k\leq K}} & yF(\sum_k n_k, \frac{\sum n_k z_k}{\sum n_k}) - \sum_k w_k n_k - \tilde{n}\frac{c}{q(\tilde{v})} - \kappa_f \\
& + \beta E_{y'|y} J(y',\{n'_k,v'_k,z'_k\}_{k\leq K+1}) \\
(\lambda_k)\ & u(w_k) + \beta[s_k U + (1-s_k)R(v'_{k+1}) = v_k]\ \forall k \leq K \\
(\omega_k) & v'_{k+1} = E_{y'|y} v'_{k+1,y'}\ \forall k \leq K \\
& n'_{k+1} = n_k(1-s_k)(1-p(v'_{k+1})) + \tilde{n}\ \forall k \leq K \\
& z'_{k+1} = min(\frac{z_k}{1-s_k},1)\ \forall k \leq K \\
& n'_0 = \tilde{n}, v'_0 = \tilde{v}, z'_0 = z_0
\end{aligned}
$$

Consider the Pareto problem

$$
\mathcal{P}(y,\{n_k,\rho_k,z_k\}) = \sup_{\{v_k\}} J(y,\{n_k,v_k,z_k\}) + \sum_k \rho_k v_k
$$

I first substitute the definition of J together with its constraints into $\mathcal{P}$:

$$
\begin{aligned}
\mathcal{P}(y,\{n_k,\rho_k,z_k\}) = \sup_{\tilde{n},\tilde{v},\{v_k,v'_k,v'_{y',k},w_k,s_k\}_{k\leq K}} & yF(\sum_k n_k, \frac{\sum n_k z_k}{\sum n_k}) - \sum_k w_k n_k - \tilde{n}\frac{c}{q(\tilde{v})} - \kappa_f \\
& + \beta E_{y'|y} J(y',\{n'_k,v'_k,z'_k\}_{k\leq K+1}) + \sum_k \rho_k v_k \\
(\lambda_k)\ & u(w_k) + \beta[s_k U + (1-s_k)R(v'_{k+1}) = v_k]\ \forall k \leq K \\
(\omega_k)\ & v'_{k+1} = E_{y'|y} v'_{k+1,y'}\ \forall k \leq K \\
& n'_{k+1} = n_k(1-s_k)(1-p(v'_{k+1})) + \tilde{n}\ \forall k \leq K \\
& z'_{k+1} = min(\frac{z_k}{1-s_k},1)\ \forall k \leq K \\
& n'_0 = \tilde{n}, v'_0 = \tilde{v}, z'_0 = z_0
\end{aligned}
$$

I now substitute in the promise-keeping constraint:

$$
\begin{aligned}
\mathcal{P}(y,\{n_k,\rho_k,z_k\}) = \sup_{\tilde{n},\tilde{v},\{v'_k,v'_{y',k},w_k,s_k\}_{k\leq K}} & yF(\sum_k n_k, \frac{\sum n_k z_k}{\sum n_k}) - \sum_k w_k n_k - \tilde{n}\frac{c}{q(\tilde{v})} - \kappa_f \\
& + \beta E_{y'|y} J(y',\{n'_k,v'_k,z'_k\}_{k\leq K+1}) + \sum_k \rho_k (u(w_k) + \beta[s_k U + (1-s_k)R(v'_{k+1})]) \\
(\omega_k)\ & v'_{k+1} = E_{y'|y} v'_{k+1,y'}\ \forall k \leq K \\
& n'_{k+1} = n_k(1-s_k)(1-p(v'_{k+1})) + \tilde{n}\ \forall k \leq K \\
& z'_{k+1} = min(\frac{z_k}{1-s_k},1)\ \forall k \leq K \\
& n'_0 = \tilde{n}, v'_0 = \tilde{v}, z'_0 = z_0
\end{aligned}
$$

I introduce the ω_k-constraints with weights β into the problem:

$$\begin{aligned}
\mathcal{P}(y,\{n_k,\rho_k,z_k\}) = \inf_{\{\omega_k\}} \sup_{\tilde{n},\tilde{v},\{v'_k,v'_{y',k},w_k,s_k\}_{k\leq K}} & yF(\sum_k n_k, \frac{\sum n_k z_k}{\sum n_k}) - \sum_k w_k n_k - \tilde{n}\frac{c}{q(\tilde{v})} - \kappa_f \\
& + \beta E_{y'|y} J(y', \{n'_k, v'_k, z'_k\}_{k\leq K+1}) + \sum_k \rho_k(u(w_k) + \beta[s_k U + (1-s_k)R(v'_{k+1})]) \\
& + \sum_k \beta\omega_k (E_{y'|y} v'_{y',k+1} - v'_{k+1}) \\
& n'_{k+1} = n_k(1-s_k)(1-p(v'_{k+1})) + \tilde{n}\ \forall k \leq K \\
& z'_{k+1} = min(\frac{z_k}{1-s_k}, 1)\ \forall k \leq K \\
& n'_0 = \tilde{n}, v'_0 = \tilde{v}, z'_0 = z_0
\end{aligned}$$

I then rearrange the value function by moving $E_{y'|y}\sum_k \beta\omega_k n'_{k+1} v'_{y',k+1}$ (additional constraints are dropped to simplify notation):

$$\begin{aligned}
\mathcal{P}(y,\{n_k,\rho_k,z_k\}) = \inf_{\{\omega_k\}} \sup_{\tilde{n},\tilde{v},\{v_k,v'_{y',k},w_k,s_k\}_{k\leq K}} & yF(\sum_k n_k, \frac{\sum n_k z_k}{\sum n_k}) - \sum_k w_k n_k - \tilde{n}\frac{c}{q(\tilde{v})} - \kappa_f \\
& + \beta E_{y'|y}[J(y', \{n'_k, v'_k, z'_k\}_{k\leq K+1}) + \sum_k \omega_k v'_{y',k+1}] \\
& \sum_k \rho_k(u(w_k) + \beta[s_k U + (1-s_k)R(v'_{k+1})]) - \sum_k \beta\omega_k v'_{k+1}
\end{aligned}$$

Lastly, I split the sup:

$$\begin{aligned}
\mathcal{P}(y,\{n_k,\rho_k,z_k\}) = \inf_{\{\omega_k\}} \sup_{\tilde{n},\tilde{v},\{v'_k,w_k,s_k\}_{k\leq K}} & yF(\sum_k n_k, \frac{\sum n_k z_k}{\sum n_k}) - \sum_k w_k n_k - \tilde{n}\frac{c}{q(\tilde{v})} - \kappa_f \\
& + \beta E_{y'|y}[\sup_{v'_{y',k+1}} J(y', \{n'_k, v'_k, z'_k\}_{k\leq K+1}) + \sum_k \omega_k v'_{y',k+1}] \\
& \sum_k \rho_k(u(w_k) + \beta[s_k U + (1-s_k)R(v'_{k+1})]) - \sum_k \beta\omega_k v'_{k+1}
\end{aligned}$$

From this, one can note that, by definition of $\mathcal{P}$

$$\sup_{v'_{y',k+1}} J(y', \{n'_k, v'_k, z'_k\}_{k\leq K+1}) + \sum_k \omega_k v'_{y',k+1} = \mathcal{P}(y', \{n'_k, \omega_k, z'_k\})$$

We thus arrive to the formulationof the problem as described at the beginning, not involving finding future state-specific promised values $v'_{y',k}$. $\square$

A.5 Block Recursivity

I introduce an assumption that would allow for a block recursive eqiulibrium under the same conditions as in Schaal (2017). Block recursivity requires an indifference condition, either on the side of the firms or on the side of the workers. Under two-sided ex-post heterogeneity, that is not immediately achievable.

Schaal (2017) shows that, in a setting similar to mine, but with transferable utility between workers and firms, which he achieves due to the risk-neutral worker utility function, firms all have the same preferences across all the submarkets that they may post vacancies in. Define the minimal hiring cost as

$$k = \min_{v}[v + \frac{c}{q_v}]$$

Due to transferable utility, the cost of employing the worker from submarket v becomes simply the value v. Thus, the optimal entry of vacancies in Schaal (2017) can be summarized by

$$\theta_v[v + \frac{c}{q_v} - k] = 0$$

Meaning that either a submarket v minimizes the hiring cost or it is closed. This condition is completely independent of the distribution of firms and workers, exactly because the one component where the firm type might come through, the cost of employing a worker from submarket v, is completely independent from the firm's state due to transferable utility.

Utility is not transferable in my model, and thus different firms may face different costs of employing a worker at some value v (for example, fixing y and z, small firms prefer high values v due to their intention to upsize). To get around that, I split the value v that the worker would get upon getting hired into two components, the sign-on wage w_v and the remaining value v_0 such that

$$u(w_v) + \beta v_0 = v$$

This additional wage payment is incurred immediately upon hiring, allowing the remaining value that the firm owes to its worker, v_0, to be completely independent of the submarket v. Essentially, from the firm's perspective, submarkets now differ not in the value that firms would owe to the workers, but in this sign-on wage. The cost minimization problem then becomes

$$k = \min_{v}[w_v + \frac{c}{q_v}]$$

This problem is now again completely independent of the firm's state, and thus the distribution of firms and workers no longer affects the tightness function q_v.Schaal (2017) shows that, in a setting similar to mine, but with transferable utility between workers and firms, which he achieves due to the risk-neutral worker utility function, firms all have the same preferences across all the submarkets that they may post vacancies in. Then setting θ_v such that

B Quantitative Appendix

B.1 Variables

Endogenous state variables

$$\{n_k\}_{k\leq K}, \{\rho_k, z_k\}_{1\leq k\leq K}$$

Code: `states: size, rho, q`

Exogenous state variables

$$y$$

Code: `z`

Control variables

$$\tilde{n}, \{\rho'_k\}_{1\leq k\leq K}, \{s_k\}_{k\leq K}$$

Code: `hiring, rho_star, sep_star`

Value function

$$\mathcal{P}(y, \{n_k, \rho_k, z_k\})$$

Code: `ERho_star, EJ_star`

B.2 Equations

Equations (1)–(21): Value function in olive (given by value and its gradient guesses from future states), states in green (given), controls in orange (given by policy guess from current states), next period exogenous states in magenta (to be summed over), next period states in blue, parameters in black.

$$0 = \rho'_{k+1} - \rho_k - \eta(E\frac{\partial P(y', \{n'_k, \rho'_k, z'_k\})}{\partial \rho'_{k+1}})E\frac{\partial \mathcal{P}(y', \{n'_k, \rho'_k, z'_k\})}{\partial n'_{k+1}} \tag{7}$$

$$\tag{8}$$

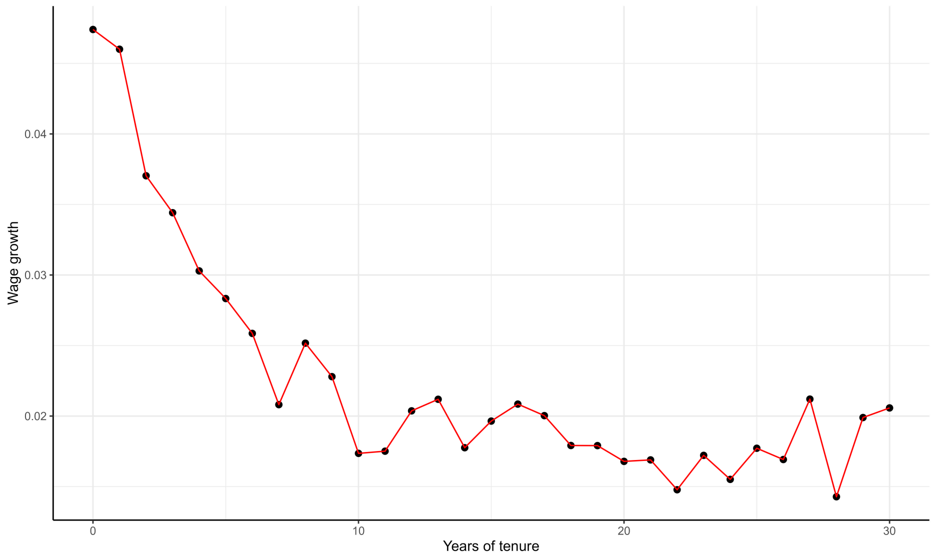
C Data Appendix

C.1 Sample Construction

C.2 Layoffs using Labor Force Survey

C.3 Wage Growth

I use the same sample to plot the log (real) wage growth across first 30 years of tenure.



The wage growth appears to flatten after about 10 years of tenure, suggesting that it is not quantitatively costly to use $K = 10$ as an approximation of the firm problem from Definition 1.