Large Firms, High Concentration, High Wages*

Guillaume NÉVO†

This version: March 23, 2026 (click here for the latest)

Abstract

How do wages and worker flows vary with labor-market concentration? Using comprehensive French administrative employer–employee data, I document that wages are 6.9% higher in high-concentration markets—contrary to much of the literature—while job-to-job transitions are lower. These patterns hold in the raw data, in market-level and panel regressions, and around large concentration increases. To interpret them, I develop a search-and-matching model with a discrete number of multi-worker firms and an endogenous operating margin. Increasing operating costs reduces the number of active firms and raises concentration, weakening workers' outside options. Yet it also induces selection: low-productivity firms exit, raising average employer quality and strengthening outside options. In the estimation, selection dominates, so mean wages rise with concentration. I characterize the planner's allocation and show that too many unproductive firms operate. Efficiency cannot generally be restored. Quantitatively, a policy taxing out low-productivity firms almost restores efficiency.

Keywords: Labor-market concentration, monopsony, firm granularity, search-and-matching, firm entry.

*I am grateful to Guido Menzio, Katarina Borovičková, and Jaroslav Borovička for their invaluable support and guidance. This project greatly improved thanks to numerous discussions with Aleksandra Alferova, Corina Boar, Pierre Cahuc, Jan Eeckhout, Niklas Engbom, Alfred Galichon, Basile Grassi, Patrick Kehoe, Francis Kramarz, Eungik Lee, Elena Manresa, Elena Pastorino, Diego Perez, Alessandra Peter, Sharon Traiberman, Venky Venkateswaran, Nicholas Zarra, and audiences at NYU, Stern, Sciences Po, Uppsala University, University of Copenhagen, Bank of Chile, and several conferences. I am especially grateful to two discussants, Aniket Baksy and Miren Azkarate-Askasua. I am also greatly indebted to Alfred Galichon and Sciences Po Paris for offering me a working space to access the data. This material is based upon work supported by the National Science Foundation under Grant No. 2242398, and by the C. V. Starr Center for Applied Economics. Access to the data has been made possible within a secure environment offered by CASD – Centre d'accès sécurisé aux données (Ref. 10.34724/CASD).

†Department of Economics, Uppsala University. Email: guillaume.nevo@nek.uu.se

Introduction

Product markets have become increasingly dominated by large firms, and the implications of rising concentration have been studied extensively. Analogous concerns have recently shifted to *labor* markets: a growing empirical literature measures concentration in local labor markets and studies how it relates to wages, mobility, and other worker outcomes. These studies emphasize that labor-market concentration increases employer market power by weakening workers' outside options, putting downward pressure on wages and job-to-job mobility. Reflecting this view, policy debates often treat labor-market concentration as a source of employer power—for example, the White House's 2021 Executive Order on Promoting Competition argues that consolidation can make it harder for workers to bargain for higher wages and better working conditions.

Yet, relatively little is known about the *general-equilibrium* forces that shape labor-market concentration—what drives it, and what it implies for workers overall. Concentration is an equilibrium outcome: it reflects employer market power, but it is also shaped by endogenous selection across firms, so more concentrated markets may be those in which employment is concentrated in more productive, higher-paying firms. Motivated by this tension, I document how labor-market concentration relates to wages and worker mobility and develop a general-equilibrium framework to interpret these patterns. I then use the framework to quantify wage markdowns, characterize efficiency, and evaluate policy counterfactuals that affect concentration and worker outcomes.

Using French administrative employer–employee data (the repeated cross-section of the universe of job spells, a representative panel, and a dataset that consolidates legal units into firms), I measure labor-market concentration and document two empirical facts. First, in France, concentration and wages *positively* co-move—opposite to much of the U.S. and Norwegian evidence (Rinz 2022; Azar, Marinescu, and Steinbaum 2022; Berger

et al. 2024). The positive relationship appears in the raw data, in market-level regressions à la Berger et al. (2024), and in panel specifications; it is robust across market definitions and concentration measures. In my preferred panel specification, moving from the 10th to the 90th concentration percentile is associated with a 6.9% increase in daily wages. In contrast, mobility patterns are in line with intuition: more concentrated markets exhibit fewer job-to-job transitions as fewer firms operate (consistent with Berger et al. 2024).

Second, to complement these estimates—which rely on relatively small year-to-year movements in concentration—I study large increases in concentration in an event-study design. In the data, these events are typically driven by rapid employment growth at a single large firm. At the market level, mean wages do not fall after such shocks (if anything, they rise slightly); at the worker level, wages increase by about 5% to 10% relative to controls for workers staying at the expanding firm, with spillovers to nearby firms and markets.

I use a structural model to interpret the mechanisms and quantify wage markdowns. The model builds on Acemoglu and Hawkins (2014) but features a *discrete* number of firms: multi-worker firms operate in a search-and-matching environment, post vacancies optimally, and set wages through Stole–Zwiebel bargaining. A novel ingredient is an *endogenous operating decision*: firms pay a per-period fixed cost to operate, which determines which firms are active and, consequently, how concentrated the market is.

The model reproduces the two main empirical findings. In the cross-section, higher operating costs reduce the number of active firms and increase concentration. Fewer firms weaken workers' outside options (a monopsony effect), but higher operating costs also select more productive firms into operation (a selection effect). Whether wages rise or fall with concentration depends on the thickness of the tail of the productivity distribution: with a thicker tail, selection raises average productivity faster than concentration raises monopsony power. I estimate the shape of the distribution to match the positive wage–concentration

relationship; in the calibrated economy, the selection effect dominates, so mean wages rise with concentration while job-to-job transitions decline.

Within a given labor market, a productivity increase at a large firm raises wages at that firm through higher output. Wages at other firms also rise because the dominant firm improves workers' outside options (even though other firms' productivity is unchanged). Markdowns increase at the dominant firm and fall at others. Job-to-job transitions rise toward the dominant firm and fall toward others.

I estimate the model by simulated method of moments and use it to quantify wage markdowns. Across markets, markdowns are modest: the average lies between 0.10 and 0.18. Decomposing wages into (i) the worker's share of output, (ii) a poaching-related term, and (iii) compensation through outside options, I find that the worker's share of output is the largest component (about 50–60% of wages). This helps explain why wages increase at more productive firms and in more concentrated labor markets.

I then use the model to study efficiency and policy counterfactuals. I characterize the first-best allocation in which a planner chooses (i) which firms operate, (ii) vacancy postings, and (iii) the allocation of workers across firms. The main inefficiency arises at the entry margin: too many low-productivity firms operate and crowd the vacancy market. In simulations, the first-best sharply reduces the number of active firms, concentrates employment among the most productive firms, and raises output by about 0.8%, at the cost of higher unemployment.

I then study how close policy can get to the first-best. With on-the-job search, efficiency generally cannot be restored because job-ladder dynamics create discrete poaching rents. Quantitatively, an entry tax does most of the work: taxing marginal low-productivity firms out of the market brings the decentralized equilibrium close to the planner's allocation.

Policy implications. These results highlight the links between productivity, concentration, and wages. More productive firms tend to be larger and pay higher wages, so

concentration and wages can rise together. Although intuitive, this channel has been largely overlooked in academic and public discussions. Markdowns exist, but they are quantitatively small relative to output sharing; as a result, large firms can raise wages precisely because they are more productive.

The counterfactuals clarify how concentration can raise wages: deterring the entry of low-productivity firms (even at the cost of higher unemployment) can increase output and wages by reallocating employment toward more productive, higher-paying firms. My results are the mirror image of Boar and Midrigan (2025) on the product-market side. The key policy question is therefore not concentration per se, but whether concentration reflects efficient selection into productive firms or distortions that weaken workers' outside options.

This perspective also sheds light on policies that subsidize low-productivity operation. In France, payroll tax exemptions targeted at low-wage jobs may keep marginal, low-productivity ("zombie") firms in operation; through the lens of the model, they work against the main efficiency margin—screening out low-productivity firms to relieve vacancy-market crowding and reallocate hiring toward more productive firms.

Related literature

My paper is most closely related to Berger et al. (2024), who study Norway using universe employer–employee data. I replicate their market-level wage regressions, and both papers build on a discrete-firm job-ladder environment.

There are also important differences. Empirically, I find the opposite sign: wages rise with concentration in France, whereas they fall in Norway. Several features of the French setting and my measurement choices help reconcile these results. I measure concentration at the business-group level, which raises measured concentration by about 20% relative to legal-unit measures but leaves the qualitative patterns unchanged. French labor markets

are also more concentrated on average (typical HHI around 0.2 versus 0.1 in Norway, even without group consolidation), suggesting that they are more often dominated by a large—and plausibly more productive—employer. As larger firms tend to pay higher wages, higher concentration can correlate positively with wages in France. If productivity differences among dominant employers within a labor market are more pronounced in France, both results can be understood: market power dominates in Norway, whereas productivity-driven selection dominates in France. I further extend the empirical analysis through worker-level panel regressions and an event-study around large concentration changes.

On the modeling side, I show how a multi-worker firm model à la Acemoglu and Hawkins (2014) can deliver an equilibrium close to that in Berger et al. (2024). A key addition is an endogenous operating margin, which generates selection into operation and reconciles the positive wage–concentration relationship. Crucially, my model can also produce the negative slope they obtain, depending on the tail of the productivity distribution. Finally, I characterize the planner's problem and study second-best taxes, which are absent from Berger et al. (2024).

More broadly, this paper contributes to the empirical literature on labor-market concentration and worker outcomes, which uses evidence from online vacancies (Azar, Marinescu, and Steinbaum 2022; Azar et al. 2020) and administrative data in the U.S. (Rinz 2022; Yeh, Macaluso, and Hershbein 2022; Autor, Patterson, and Van Reenen 2023), France (Bassanini, Batut, and Caroli 2023; Marinescu, Ouss, and Pape 2021), and other European countries (Bassanini et al. 2024). Most of this literature finds a negative wage–concentration relationship; differences may reflect the outcome studied (often new hires) and the estimation methods (such as IV strategies).

Using the same administrative data, Marinescu, Ouss, and Pape (2021) provides a useful benchmark. They isolate a market-power component by combining labor- and product-

market concentration measures, defining labor-market concentration using employment shares among *new hires*, and conditioning on firm productivity (e.g., firm fixed effects and/or value added per worker). When I use employment-share HHIs and add worker and firm fixed effects, the wage–concentration relationship moves toward theirs (becoming much smaller and sometimes insignificant or negative); similar considerations apply to Bassanini, Batut, and Caroli (2023) and Bassanini et al. (2024). Two differences remain central. First, I measure concentration using wage-bill shares, which better matches the model's notion of outside option. In the model, outside options average match surpluses across potential employers, scaled by each firm's vacancy share; wage-bill shares incorporate both employment and wages and better proxy for firms' contributions to workers' outside options than employment shares alone. Second, their baseline specifications condition on firm productivity (through product-market concentration, firm fixed effects, and/or value added per worker). This is well-suited to isolating a market-power component holding productivity fixed, but it abstracts from the productivity-driven selection component of concentration that is central to my interpretation. Overall, these approaches target different objects in the same data: theirs isolates the wage effect of market power holding productivity fixed, whereas mine interprets concentration as an equilibrium outcome; accordingly, the results are not inconsistent. Appendix D.5 illustrates this distinction in simulated panel data from the model by showing how conditioning on productivity (or using firm fixed effects) removes the output-sharing/selection component of wages and mechanically attenuates the concentration coefficient, bringing the regression closer to a "market power" object.

On the theory side, this paper relates to models of search frictions and monopsony power (Burdett and Mortensen 1998; Manning 2003; Postel–Vinay and Robin 2002) and to recent work emphasizing firm granularity (Jarosch, Nimczik, and Sorkin 2024; Berger, Herkenhoff, and Mongey 2022). A novel feature is a discrete-firm equilibrium with endogenous operation

and vacancy creation, which generates firm selection (in the spirit of Hopenhayn 1992 and Melitz 2003) and an endogenous concentration distribution. Related mechanisms arise in Lindenlaub, Oh, and Peters (2024), where sorting of productive firms across locations can jointly deliver higher wages and a lower labor share (higher markdowns).

This paper is also related to Bilal (2023), who develops a general-equilibrium framework in which spatial differences in unemployment are driven primarily by employer heterogeneity and equilibrium allocation across local labor markets. While I abstract from endogenous location choice, my analysis similarly treats local labor-market structure as an equilibrium object shaped by firm heterogeneity and selection; in both settings, this motivates policy counterfactuals that operate on the firm entry margin.

Finally, my paper relates to the literature on efficiency in frictional labor markets (Hosios 1990; Mortensen and Pissarides 1994; Fukui and Mukoyama 2025). Relative to this work, I analyze a planner's problem with a discrete number of firms and an endogenous operation margin, which adds new margins to the planner's allocation. I then study second-best implementations through an entry tax and a production tax. The key inefficiency at the operation margin is closely related to Acemoglu (2001): when heterogeneous firms/jobs coexist, low-productivity ("bad") firms do not internalize that their entry and hiring reduce the availability of workers and vacancies for high-productivity ("good") firms, so the decentralized equilibrium features too much low-productivity activity.

1 Data

1.1 Datasets and data selection

I use three French administrative employer–employee datasets collected by INSEE: (i) *Base Tous Salariés – fichier Postes* (BTS), a repeated cross-section covering the universe of jobs;

(ii) *Panel Tous Salariés* (PTS), a 10% matched employer–employee panel; and (iii) *Contour des Entreprises Profilées* (CEP), which links legal units belonging to the same business group using financial and ownership information. These datasets have been used extensively in prior work (e.g. Abowd, Kramarz, and Margolis 1999; Bilal 2023). I focus on the years 2009–2019.

BTS provides a repeated cross-section of all jobs at the establishment level, with demographic, geographic, and detailed occupational information. Because worker identifiers are reshuffled between waves, individuals cannot be followed beyond a two-year window. PTS is a 10% matched employer–employee panel drawn from individuals born in October. Finally, I use CEP for employer identification—I group legal units belonging to the same business group within each labor market into a single firm.

I restrict the analysis to private-sector, full-time jobs for workers aged 20–60 in mainland France and drop jobs paying less than one monthly minimum wage over the year. Jobs are assigned to commuting zones using INSEE's 2020 definition, and occupations are harmonized using the PCS–ESE (2003) classification. Wages are expressed in 2019 euros. For each worker-year, I keep a single observation—the highest-paying job, which I refer to as the main spell. Labor market status and transitions—stayers, job-to-job (J2J), job-to-unemployment (J2U), and unemployment-to-job (U2J)—are defined from non-employment days between consecutive main spells; full classification rules are provided in Appendix A.

1.2 Labor market definition and concentration measures

I define a labor market as occupation (3-digit) $\times$ commuting zone $\times$ year. A key requirement for this definition to be meaningful is that labor markets are sufficiently self-contained, so that it captures the notion of workers' relevant outside options. Figure 1 plots the probability of staying in the same labor market between t and $t+1$ against the percentile of the

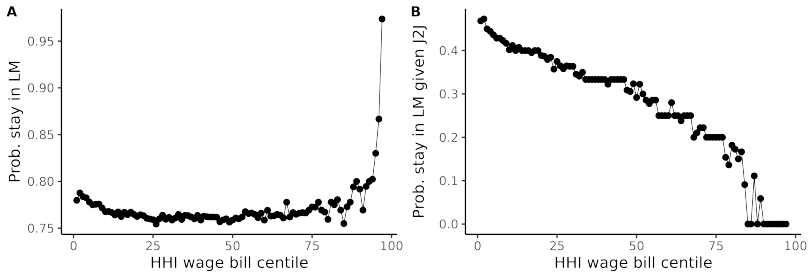


Figure 1: Probability of staying in the same labor market in two consecutive years

concentration measure in year t. Panel A shows the unconditional probability, while Panel B conditions on J2J transitions. The median unconditional and conditional self-flow rates are 76% and 33%, in line with Berger et al. (2024). Conditional self-flows decline with concentration, but this decline is concentrated in a small number of highly specialized markets with few workers.[1]

I measure concentration using the Herfindahl–Hirschman Index (HHI) based on firms' wage-bill shares within a labor market. For each occupation $\times$ commuting zone $\times$ year cell, I count the number of firms employing workers in their main spell, N_{oct}^{firms}, and compute each firm's wage-bill share s_i^{wb}. The HHI is then:

$$\text{HHI}_{oct}^{wb} = \sum_{i=1}^{N_{oct}^{firms}} (s_i^{wb})^2, \qquad s_i^{wb} = \frac{\text{wage bill}_i}{\sum_k \text{wage bill}_k}.$$

The HHI is a measure between 0 (many firms with similar wage-bill—low concentration) and 1 (single firm—high concentration). As a benchmark, if all firms have equal shares, then HHI $= 1/N$.

[1]Alternative labor-market definitions using workers' mobility patterns (e.g., Schmutte 2014; Nimczik 2020) or clustering methods (e.g., Berger et al. 2024), have been proposed but do not significantly improve the self-flow rates. As a robustness check, I verify that my results are robust to using 2-digit occupations.

The HHI is standard in the product-market literature (e.g., Edmond, Midrigan, and Xu 2015; Burstein, Carvalho, and Grassi 2025) and increasingly used in labor-market applications (e.g., Deb et al. 2024; Berger, Herkenhoff, and Mongey 2022; Berger et al. 2024). Appendix B reports robustness to measures based on firm employment shares.

2 Concentration, wages, and mobility in France

This section documents how concentration co-moves with labor-market outcomes using market-level regressions (as in Berger et al. 2024) and panel regressions. I then study large within-market increases in concentration in an event-study setting. Additional descriptive evidence and robustness checks are reported in the data appendices (Appendices A–C).

2.1 Labor-market concentration and labor-market outcomes

Figure 2 summarizes labor-market concentration in France, measured by HHI^{wb}. The figure reports the median concentration, median mean wage and median J2J probability across labor markets within the same HHI percentile.[2]

Using the U.S. Department of Justice threshold of 0.15 for a moderately concentrated market as a benchmark, about two-thirds of labor markets fall below it as seen in Panel A. Labor markets in low-HHI percentiles are much larger: a market in the 10th percentile has on average about 150 firms and 400 workers, while a market around $\text{HHI} = 0.15$ has about 20 firms and 50 workers; by the 90th percentile, these numbers fall to about 4 firms and 13 workers. About 91% of workers and 94% of firms fall under the 0.15 threshold.

Workers are disproportionately located in large commuting zones (especially Paris),

[2]There are around 400k markets (120 3-digit occupations $\times$ 300 commuting zones $\times$ 12 years). A percentile contains approximately 4k markets.

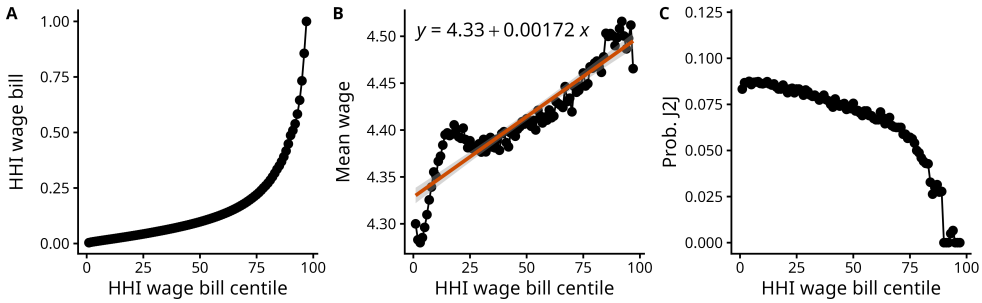


Figure 2: Concentration level, mean wage and J2J probability across labor markets by HHI^{wb} percentile

which mechanically have lower concentration. This creates selection issues. To alleviate this concern, I focus on specifications that exploit within-market changes over time.

In the raw data, Figure 2 shows that wages increase with concentration (Panel B) and job-to-job mobility declines with concentration (Panel C). Between the lowest and highest HHI percentiles, real wages increase by 17%, while the job-to-job transition probability declines steadily (from 8.5% in the lowest percentile to 0% in the highest percentile). Additional descriptive patterns (age composition, gender composition, and other transition rates) are reported in Appendix A.

2.2 Removing compositional effects

I now control for compositional effects using (i) market-level regressions in the spirit of Berger et al. (2024), which regress mean log daily wage and J2J probability in a labor market in a given year on concentration, and (ii) worker-level panel regressions, which regress log daily wage and transition indicators on concentration.

In both cases, my preferred specification includes occupation $\times$ commuting-zone fixed effects. These labor-market fixed effects absorb time-invariant market-specific factors

	Mean log daily wage	
	(i)	(ii)
HHI^{wb}	0.0820^{***}	0.0500^{***}
	(0.0058)	(0.0060)
Occ.×CZ FEs	✓	✓
Add. controls		✓
Observations	407K	407K
R^2	0.905	0.909
Within R^2	0.006	0.045

Clustered (LM) standard errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 1: Market-level regressions: mean log daily wage on concentration

(including average composition), so the coefficient on concentration is identified from within-market changes across years—for instance due to hiring and separations, or firm entry and exit—and how it is associated to changes in the dependent variable. My baseline wage market-level regression is:

$$\text{mean log daily wage}_{oct} = \nu_{oc} + \beta\,\text{HHI}_{oct} + \gamma X_{oct} + \epsilon_{oct}$$

where ν is a labor-market fixed effect and X is a vector of controls mirroring Berger et al. (2024) (fraction of men, quartiles of age composition, and the ratio of workers to firms). Standard errors are clustered at the labor-market level. J2J probability regression is similar.

Table 1 reports the market-level estimates. Controlling for labor-market fixed effects, higher concentration is associated with higher mean log daily wages, with strongly significant coefficients around 0.05–0.08 depending on controls. To interpret magnitudes, Figure 2 shows that markets around the 10th and 90th HHI percentiles have HHI ≈ 0.02 and HHI ≈ 0.49. Hence, multiplying the coefficient by ≈ 0.5 implies an approximate 2.5–4.1%

	Log daily wage (i)	Stay at firm (ii)	J2J (iii)	J2U (iv)	U2J (v)
HHI^{wb}	0.1380***	0.0237***	-0.0314***	-0.0220***	-0.0223***
	(0.0120)	(0.0059)	(0.0082)	(0.0045)	(0.0045)
Sex & Occ.×CZ FEs	✓	✓	✓	✓	✓
Observations	8.7M	7.2M	7.2M	7.2M	7.2M
R^2	0.560	0.108	0.025	0.094	0.068
Within R^2	0.127	0.054	0.011	0.042	0.029

Clustered (firm) standard errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 2: Panel data regressions: log daily wage and transitions

difference in mean log daily wage between these two percentiles. While much smaller than the raw correlation ((0.90-0.10)×17%), it remains economically meaningful. Robustness checks yield consistent results (Appendix B.1).

I now turn to the panel data, which allow me to control for more worker characteristics. The empirical specification is:

$$y_{it} = \nu_{o(i)c(i)} + \beta \, \text{HHI}^{wb}_{o(i)c(i)t} + \gamma X_{it} + \epsilon_{it}$$

where y is either log daily wage or mobility indicators (stay, J2J, J2U, and U2J). I control for a third-degree age polynomial, sex and labor-market fixed effects, and add tenure and work-experience controls in extended specifications in the appendix. Table 2 reports the results; additional specifications are in Appendix B.2. Standard errors are clustered at the firm level.

Moving from the 10th to the 90th HHI percentile is associated with about a 6.9% increase in daily wages and a 1.6 percentage-point decline in job-to-job mobility. Workers are more

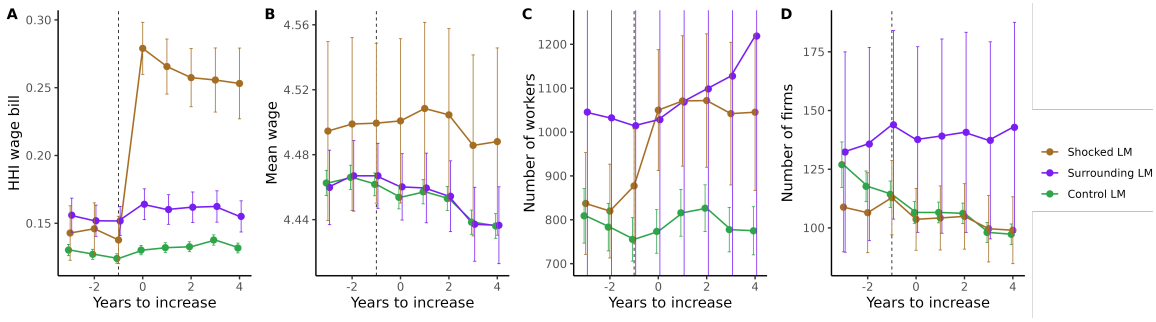


Figure 3: Large HHI increase on market-level outcomes

likely to stay at their current firm, and less likely to transition to or from unemployment. This is consistent with fewer outside options in more concentrated markets.

2.3 Focusing on large changes in concentration

I now focus on large changes in concentration within a market. Such shocks are policy-relevant: major expansions, closures, and mergers attract attention and can affect many workers. They also complement the previous within-market estimates, which rely on relatively small year-to-year movements in concentration.[3]

I construct an event-study sample of large labor markets experiencing large within-market changes in HHI^{wb}. I compare them to control markets in the same occupation but in commuting zones that are not geographically adjacent to the shocked markets. I also track markets adjacent to the shock to assess spillovers. More details on the event-study are available in Appendix C.

Large increases in HHI^{wb} can reflect either the expansion of a large firm or the exit

[3]Concentration is highly persistent within markets: in an AR(1) regression with no intercept for HHI over time, the median coefficient across markets with at least 10 workers is 0.975.

of competing firms. Figure 3 suggests the former: following the shock (year 0), the HHI rises sharply in shocked markets (Panel A) while remaining flat in surrounding and control markets; employment increases strongly (Panel C) and the number of firms falls only slightly (Panel D). Appendix C shows that the large increase is driven by the rapid expansion of a single firm, suggesting an increase in demand or productivity at the firm level.[4] Yet, despite the increase in concentration, mean wages do not fall and, if anything, rise relative to control markets (Panel B).

The appendix further reports a worker-level matched event-study that matches pre-shock workers to comparable controls. It suggests wage gains following the shock, although a pre-trend is present for workers at the expanding firm; mobility patterns are consistent with workers reallocating toward the expanding firm.

I also examined labor markets experiencing large HHI decreases. These cases are typically driven by contractions of dominant firms. Average wages tend to decrease in shocked labor markets compared to control markets. At the firm-level though, results are difficult to interpret causally (e.g. due to selection of workers to lay-off—last-in-first-out restrictions—mechanically tilting the wage distribution at the shocked firm to the right).

Taken together, I document two main empirical correlations: (i) in the cross-section, more concentrated labor markets exhibit higher wages and lower job-to-job mobility; and (ii) in the event-study setting, large within-market increases in concentration are, if anything, followed by wage increases.

[4]Cestone et al. (2024) study one concrete source of such expansion opportunities—the exit of a large competitor—and show that firms belonging to the same business group expand by redeploying workers across group units through internal labor markets, reallocating labor from low to high marginal product of labor firms.

3 A model of the labor market with endogenous entry

I adopt a structural approach that moves beyond empirical correlations to quantify the effects of firms' market power on workers. A search-and-matching model with multi-worker firms and bargaining, following Acemoglu and Hawkins (2014), provides a natural framework for studying how wages, labor market transitions, and unemployment are affected by labor market concentration. A key departure from standard models is firm granularity (e.g., Jarosch, Nimczik, and Sorkin 2024).

Consider a single, closed labor market with a discrete number of firms and a unit mass of workers. Firms and workers are infinitely lived, risk-neutral, and discount the future at a common rate ρ. Potential entrants draw productivity and decide whether to operate in the market (paying the per-period operating cost c_f). I first abstract from the operating decision to highlight the model's mechanisms for a given number of firms and return to firms' operating decisions at the end of this section. Conditional on operating, firms choose vacancy postings (and hence firm size) optimally, trading off vacancy-posting costs against the expected gains from hiring additional workers. Firms internalize the effect of v_i on their own arrival rates (including the congestion effect) but treat the worker distribution and other firms' choices as given.

On the worker side, individuals are *ex-ante* identical and equally productive; they search both on- and off-the-job. Search is random, and wages are determined by Stole–Zwiebel bargaining (Stole and Zwiebel 1996) over the marginal surplus of the employment relationship, as in Acemoglu and Hawkins (2014). I focus on a stationary equilibrium with complete information. Firms produce under constant returns to scale using labor as the sole input.

These assumptions are crucial because, as in a standard search-and-matching model, they allow me to solve the firm's problem at the match (bilateral) level. Combined with

optimal vacancy posting, the framework delivers an endogenous distribution of workers across a discrete set of firms and a consistent firm operating decision.

3.1 Contact rates between firms and workers

Firm granularity requires an explicit model of contact rates between workers and firms; I follow Berger et al. (2024). Let λ^j (resp. λ_0^j) denote the offer arrival rate to an employed worker (resp. an unemployed worker) from firm j. Let v_j be the mass of vacancies posted by firm j, and let a^j be the mass of workers who apply to j (regardless of whether they ultimately receive an offer).[5] The contact rate with firm j is:

$$\lambda^j = \Phi \cdot \xi \cdot \frac{v_j}{v} \cdot m(v_j, a^j) \cdot \frac{1}{a^j}$$

Workers search with probability Φ (with on-the-job search efficiency ξ). Conditional on searching, they meet a firm in proportion to its vacancy share v_j/v, where $v = \sum_j v_j$. Given v_j vacancies and a^j applicants, the matching function m delivers $m(v_j, a^j)$ contacts at firm j; uniform sampling among applicants then implies an offer probability $m(v_j, a^j)/a^j$. For unemployed workers, the contact rate λ_0^j is the same expression with unemployed search efficiency normalized to one.

Let n_k denote the mass of workers employed at firm k, and let u denote the mass of unemployed workers. Total applicants to firm j are the sum of applicants from all firms and from unemployment, $a^j = \sum_k a_k^j + a_0^j$, where for employed workers: $a_k^j = n_k \cdot \Phi \cdot \xi \cdot \frac{v_j}{v}$, since each worker at k meets a vacancy with probability $\Phi\xi$, and—conditional on meeting a vacancy—meets a vacancy from j with probability v_j/v. Similarly, $a_0^j = u \cdot \Phi \cdot \frac{v_j}{v}$. Finally,

[5] Conditional on employment status, meeting probabilities do not depend on the worker's current firm, so I suppress i in the notation.

using $\sum_k n_k = 1 - u$, I can write a^j as $a^j = \Phi \cdot (u + (1-u)\xi) \cdot \frac{v_j}{v}$. Substituting a^j into the expressions for λ^j and λ_0^j yields:

$$\lambda^j = \frac{\xi}{u + (1-u)\xi} \cdot m(v_j, a^j), \qquad \lambda_0^j = \frac{1}{u + (1-u)\xi} \cdot m(v_j, a^j)$$

Define the contact rate between a vacancy posted by firm j and a worker from firm i as:

$$\nu_i^j = \frac{m(v_j, a^j)}{v_j} \cdot \frac{a_i^j}{a^j}.$$

Firm j posts v_j vacancies and receives a^j applicants, generating $m(v_j, a^j)$ contacts; thus, each vacancy meets an applicant at rate $m(v_j, a^j)/v_j$. Conditional on meeting an applicant, the applicant comes from firm i with probability a_i^j/a^j. Substituting for a_i^j and a^j yields:

$$\nu_i^j = \frac{m(v_j, a^j)}{v_j} \cdot \frac{n_i \xi}{u + (1-u)\xi}, \qquad \nu_0^j = \frac{m(v_j, a^j)}{v_j} \cdot \frac{u}{u + (1-u)\xi}$$

By construction, the mass of workers from i who meet firm j equals the mass of vacancies posted by j that meet a worker from i, so $n_i \lambda^j = v_j \nu_i^j$.

3.2 Value functions, vacancy posting and bargaining

I now introduce the value functions, which satisfy Hamilton–Jacobi–Bellman (HJB) equations. Let U denote the value of unemployment, V_i the value of employment at firm i, and J_i the value of firm i. Because firms produce under constant returns to scale with labor as the sole input, output is $z_i n_i$ and firm value depends on employment, $J_i = J_i(n_i)$. Consequently, and as in Acemoglu and Hawkins (2014), V_i also depends on n_i.

Upon meeting a firm, a worker moves if and only if the value at the new firm exceeds her current value. I index firms by productivity $z_1 \leq \cdots \leq z_N$ and focus on equilibria where

the ranking of workers' values $\{V_i\}_{i=1}^N$ coincides with the ranking of productivities $\{z_i\}_{i=1}^N$.

Employment at firm i evolves due to exogenous job destruction at rate δ (workers return to unemployment), poaching by higher-ranked firms, and hiring from unemployment or from lower-ranked firms. Firms are myopic with respect to these general-equilibrium feedbacks. They internalize the effect of v_i on λ_0^i through (i) their vacancy share v_i/v holding v fixed and (ii) the total mass of vacancies v, which affects congestion. They do not internalize effects on the worker distribution, unemployment, or other firms' vacancy choices. For $j \neq i$, $\frac{\partial v_j}{\partial n_i} = 0$, $\frac{\partial n_j}{\partial n_i} = 0$, and $\frac{\partial u}{\partial n_i} = 0$.

Let v_{-i} denote other firms' vacancy postings, taken as given by i. Then $J_i(n_i)$ satisfies

$$\rho J_i(n_i) = -c_f + (z_i - w_i(n_i))n_i + \max_{v_i \geq 0} \Bigg\{ -c(v_i) + \Bigg(-\Big(\xi \sum_{j=1}^{N} \lambda_0^j(v_j, v_{-j}) \mathbb{1}_{\{V_j > V_i\}} + \delta\Big) n_i + \lambda_0^i(v_i, v_{-i}) \Big(u + \xi \sum_{j=1}^{N} n_j \mathbb{1}_{\{V_i > V_j\}}\Big) \Bigg) \frac{\partial J_i}{\partial n}\Big|_{n_i} \Bigg\}$$

where c_f is an operating cost. This term is central to the paper: it disciplines firms' operating decisions and thereby endogenizes market structure. I take the set of operating firms as given and return to the operating decision when I close the model. Production is $z_i n_i$ and the firm pays a common wage $w_i(n_i)$, determined by bargaining as described below. Vacancy posting trades off costs $c(v_i)$ against the marginal value of expanding the workforce, $\partial J_i(n_i)/\partial n$, accounting for separations (job destruction and poaching to higher-ranked firms) and hires (from unemployment and from lower-ranked firms).

The optimal vacancy posting satisfies a first-order condition. With continuous renegotiation, the current wage does not affect vacancy choice; v_i matters only through future employment. Assume a Cobb–Douglas matching function, $m(v, a) = v^{1-\theta} a^{\theta}$,[6] and take

[6] I also use the Cobb–Douglas matching function because of the efficiency properties related to the Hosios condition discussed in Section 5.

other firms' vacancy postings v_{-i} as given. The optimal vacancy posting solves:

$$c'(v_i) = \frac{\Lambda}{v}\left(\xi\theta\sum_{j>i}^{N}\frac{v_j}{v}\,n_i + \left(1-\theta\frac{v_i}{v}\right)\left(u+\xi\sum_{j<i}^{N}n_j\right)\right)\frac{\partial J_i(n_i)}{\partial n}, \tag{1}$$

where $\Lambda = \Phi^{\theta}\left(\frac{v}{u+(1-u)\xi}\right)^{1-\theta}$ and I use the ranking assumption stated above. In particular, $\lambda_0^i = \Lambda\frac{v_i}{v}$, so $\Lambda = \sum_{i=1}^{N}\lambda_0^i$ is the aggregate unemployed contact rate. If $\partial J_i/\partial n > 0$ and the firm enters, then $v_i > 0$. Moreover, for a given worker distribution and taking v_{-i} as given, the FOC admits a unique solution for v_i.[7] Equilibrium vacancy postings therefore solve a fixed point in $\{v_i\}_{i=1}^{N}$.

The unemployed worker's HJB is

$$\rho U = b + \sum_{j=1}^{N}\lambda_0^j\left(V_j - U\right)\mathbb{1}_{\{V_j>U\}}.$$

Unemployed workers receive benefits b and accept an offer from firm j whenever $V_j > U$. I assume that neither firms nor workers internalize that U depends on the equilibrium distribution of workers through λ^i.

The employed worker's HJB at firm i is

$$\rho V_i(n_i) = w_i(n_i) + \sum_{j=1}^{N}\lambda^j\left(V_j - V_i\right)\mathbb{1}_{\{V_j>V_i\}} + \delta\left(U - V_i\right) + \dot{n}_i\frac{\partial V_i(n_i)}{\partial n},$$

where $\dot{n}_i$ is the law of motion of the workforce at i implied by the firm's optimal vacancy policy v_i. Because bargaining is over the marginal worker, w_i can depend on n_i, so V_i

[7] Provided that $c'(\cdot)$ is strictly increasing with $\lim_{v\downarrow 0}c'(v) = 0$ and $\lim_{v\uparrow\infty}c'(v) = \infty$. This condition is satisfied by the usual convex vacancy-posting costs, e.g. $c(v) = \frac{1}{1+\gamma}(\varphi_v v)^{1+\gamma}$ which I use when estimating the model.

depends on n_i and on how n_i evolves over time.

Wages are determined by Nash bargaining over the incremental surplus generated by the marginal worker (Stole and Zwiebel 1996), with unemployment as the worker's outside option. This implies

$$\alpha \frac{\partial J_i(n_i)}{\partial n} = (1-\alpha)\left(V_i(n_i) - U\right).$$

Here α is the worker's bargaining power. Each worker increases the firm's surplus by $\frac{\partial J_i(n_i)}{\partial n}$, and each worker receives a flow value $V_i(n_i) - U$ by working at i. The Stole–Zwiebel bargaining protocol equalizes both.

Combining the bargaining condition with the HJB equations delivers $w_i(n_i)$. In general, w_i solves a differential equation in n_i, which depends on the distribution of the workers across firms. Under constant returns to scale, this equation is degenerate, so w_i does not depend on n_i. Appendix D.1 shows that

$$w_i = \alpha z_i + (1-\alpha)b + (1-\alpha)\Lambda \sum_{j=1}^{N} \left(V_j(n_j) - U\right) \mathbb{1}_{\{V_j > U\}} \frac{v_j}{v} - \alpha\Lambda\xi \sum_{j>i}^{N} \frac{\partial J_j(n_j)}{\partial n} \frac{v_j}{v},$$

Define the marginal surplus at firm i as: $S_i(n_i) \equiv \frac{\partial J_i(n_i)}{\partial n} + V_i(n_i) - U$. I show that S_i is independent of n_i and solves

$$(\rho+\delta)S_i = z_i - b - \alpha\Lambda \sum_{j=1}^{N} S_j \mathbb{1}_{\{S_j > 0\}} \frac{v_j}{v} + \Lambda\xi \sum_{j>i}^{N} \left(\alpha S_j - S_i\right) \frac{v_j}{v}. \tag{2}$$

Using the bargaining condition to eliminate $\partial J_i/\partial n$ and $V_i - U$, the wage can be written in terms of $\{S_j\}_{j=1}^N$:

$$w_i = \alpha z_i + (1-\alpha)b + (1-\alpha)\alpha\Lambda \sum_{j=1}^{N} S_j \mathbb{1}_{\{S_j > 0\}} \frac{v_j}{v} - (1-\alpha)\alpha\Lambda\xi \left(\sum_{j>i}^{N} S_j \frac{v_j}{v} \right). \tag{3}$$

Given $\{v_j\}_{j=1}^N$, the wage–surplus system is equivalent to a standard on-the-job search-and-matching model with bargaining from unemployment and a discrete number of firms, close to Berger et al. (2024).

The wage equation (3) at firm i has three components: the worker's share of output, αz_i; a pre-emptive markdown reflecting that employed workers can be poached, $-(1-\alpha)\alpha\Lambda\xi\sum_{j>i}^N S_j\frac{v_j}{v}$; and compensation for giving up unemployment, the worker's outside option, through benefits b and the option value of offers while unemployed, $\alpha\Lambda\sum_{j=1}^N S_j\frac{v_j}{v}$. The second term is the least intuitive. Because employed workers can be poached by higher-ranked firms, their continuation value includes the option to move to $j > i$ and obtain αS_j; this option value is reflected in the last term of (2). Anticipating this, the current firm extracts a share $1-\alpha$ of that option value in bargaining, which shows up as a wage markdown. Workers' outside option dominates the poaching penalty.

The role of concentration is already apparent in the wage equation through this outside option: when more firms operate, offer arrival rates rise, yet if marginal operating firms are lower-quality on average, the expected surplus per offer falls. Thus, changes in market concentration shift the outside option and, in turn, wages.

3.3 Markdowns

The model parallels the product-market literature on markups. Here I study wage markdowns, defined as the wedge between the wage and the marginal product of labor (equal to z_i). Define the markdown at firm i, η_i, by: $w_i = (1-\eta_i)\cdot z_i$. The markdown satisfies

$$\eta_i = (1-\alpha)\left(1 - \frac{b + \alpha\Lambda\sum_{j=1}^N S_j\frac{v_j}{v}\left(1-\xi\,\mathbb{1}_{\{j>i\}}\right)}{z_i}\right). \tag{4}$$

When workers have full bargaining power (i.e. $\alpha = 1$), firms cannot impose markdowns.

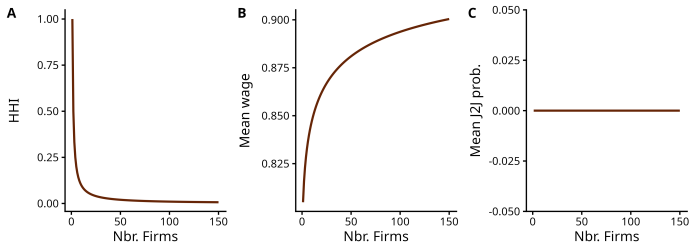


Figure 4: Mechanism of the model

Moreover, lower worker bargaining power implies a larger markdown. Although markdowns are bounded above by $1 - \alpha$, I show below that they are well below this bound, so firms capture only a small share of output.

To isolate search frictions and concentration, suppose all firms are equally productive.[8] Figure 4 shows that more firms decrease concentration (panel A), and raise wages (panel B) by improving workers' outside options. With identical productivities, poaching is absent and therefore job-to-job transitions are zero (panel C). Thus the benchmark implies a negative relationship between concentration and wages: increasing N lowers concentration and raises wages via the outside options, as in Berger et al. (2024). Introducing productivity differences would generate the negative correlation between concentration and J2J probability, but would not restore the positive correlation between concentration and wages.

3.4 Firms' operating decisions

I close the model by investigating firms' operating decisions. So far, I have abstracted from the profitability condition $J_i > 0$. When $c_f = 0$, any firm that posts a positive mass of vacancies finds it profitable to operate. With an operating cost $c_f > 0$, operation is

[8] It is straightforward to re-write the system of equations when $z_i = z$ at all firms, and there is no-poaching as there is no job-ladder. Equations are therefore omitted.

no longer guaranteed. This cost is crucial for generating a positive relationship between concentration and mean wages: raising c_f restricts operation to more productive firms and increases concentration. The effect on mean wages is ambiguous *ex ante*—selection raises average productivity, while reduced competition can worsen workers' outside options.[9] In the estimated model, most of a worker's wage reflects her share of output, so the selection channel dominates: higher concentration through restricting operation to more productive firms raises mean wages.

I characterize firms' operating decisions at the stationary equilibrium. In particular, when deciding whether to operate, firm i evaluates its stationary value J_i^*—the firm HJB evaluated at $\dot{n}_i = 0$—given the set of firms active in the market. At a stationary allocation, J_i^* satisfies

$$\rho J_i^* = -c_f + (z_i - w_i)\, n_i - c\big(v_i^*\big), \tag{5}$$

where v_i^* is the optimal steady-state vacancy posting. Firm i operates if and only if $J_i^* > 0$.[10] An equilibrium is therefore defined as:

Definition 1 (Equilibrium)**.** *A stationary equilibrium with N firms is a collection of objects $(\{z_i\}_{i=1}^N, \{S_i\}_{i=1}^N, \{J_i^*\}_{i=1}^N, \{v_i\}_{i=1}^N, \Lambda, \{n_i\}_{i=1}^N, u, \{w_i\}_{i=1}^N)$ such that, given $\{z_i\}_{i=1}^N$:*

1. *$\{S_i\}$ solves equation 2;*
2. *$\{v_i\}$ is the unique Nash equilibrium defined by equation 1;*
3. *Λ satisfies $\Lambda = \Phi^\theta \left(\frac{v}{u+(1-u)\xi}\right)^{1-\theta}$;*

[9] Because transition rates adjust, the net effect on workers' outside options is not obvious: fewer firms reduce congestion and increase the probability of meeting a productive firm, but they also reduce total vacancies and may lower overall offer arrival rates.

[10] Allowing for dynamic entry and exit, or for a transition path from operation to the stationary allocation, would require tracking the distribution of workers across firms as a state variable and would substantially complicate the analysis.

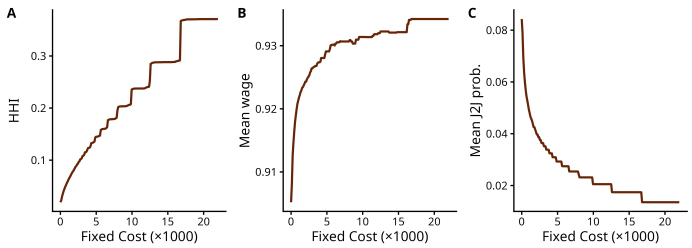


Figure 5: Operating cost and labor market outcomes

4. $\{n_i\}$ *and* u *are the unique stationary distributions implied by the transition rates and poaching rules;*

5. $\{w_i\}$ *satisfies equation 3;*

6. *Entry satisfies* $J_i^* > 0$ *for all operating firms, where* J_i^* *solves equation 5.*

The equilibrium is not unique because profits depend on which firms operate. In simulations, this indeterminacy matters only for firms at the fringe and has little quantitative impact. I therefore select the equilibrium in which operation follows the productivity ranking: when evaluating whether to operate, firm i computes J_i^* assuming that all firms $j > i$ operate.

How does the operating cost affect labor market outcomes? I fix parameters at their estimated values (Section 4) and simulate labor markets under different values of c_f. Appendix D.3 describes the algorithm. Figure 5 reports average outcomes across simulations. The operating cost is sufficient to replicate the main correlations observed in the raw data: raising c_f increases concentration (panel A) and the mean wage (panel B), while reducing job-to-job flows (panel C). As we will see, mean markdowns increase as fewer firms compete for workers, yet remain relatively small, allowing the selection channel to dominate.

Appendix D.2 shows that the model qualitatively replicates the event-study patterns in

the data. I perform a first-order perturbation of the steady-state equilibrium by introducing a productivity increase at the dominant firm and solve the resulting linear system. Relative to the pre-shock steady state, wages rise at the dominant firm (via the worker's share of output) and at other firms (via the outside option), and worker transitions also reallocate toward the dominant firm through its higher vacancy posting. When choosing a productivity increase that matches the wage increase observed at the largest firm in the event-study, the model generates an increase in employment-weighted wages of around 0.1% and in the J2J probability of around 0.8% at other firms—qualitatively in line with, but an order of magnitude lower than observed in the data.

4 Estimation

I estimate the model using a method of simulated moments. Because the model features a finite number of firms with heterogeneous productivity draws, I simulate many independent labor markets and match moments computed from the stationary equilibrium.

Each market is solved using the finite-difference approach in Achdou et al. (2022) to compute the HJBs and the implied stationary worker distribution, combined with a fixed-point iteration for vacancy posting; I then verify that the set of operating firms is consistent with profitability. Appendix D.3 sketches the solution algorithm. Because wages and transition rates are available in closed form from the equilibrium objects, I do not simulate individual worker histories.

4.1 Estimation results

Table 3 summarizes the parameters that I hold fixed throughout the estimation. I set the discount rate to 5% annually and use a standard matching elasticity of 0.5. As a benchmark,

Name	Parameters	Value	Target/Source
Discount rate	ρ	0.05	5% annual interest rate
Matching function	θ	0.5	Standard
Worker's bargaining power	α	0.5	Hosios condition
Job destruction rate	δ	0.074	BTS
Unemployment benefits	b	0.6	Standard
Number of firms drawing productivity	N	154	BTS

Table 3: Fixed parameters

I set worker bargaining power to satisfy the Hosios condition, which is linked to the model's efficiency properties. I verify in Appendix E that my main results remain valid when $\theta = 0.35$ or $\theta = 0.65$, and adjust α accordingly to satisfy the Hosios condition. I discipline the job destruction rate using separation rates observed in the low-concentration labor market in the BTS (taken to be the median market in the 10th percentile of the concentration distribution), and set unemployment benefits to $b = 0.6$, in line with Menzio and Shi (2010) or Fukui and Mukoyama (2025). Finally, the number of firms drawing productivity matches the number of firms operating in a low-concentration labor market.

I estimate the remaining parameters by matching moments between the data and the model. I rank my data markets based on their concentration level. For each percentile of concentration, I compute the median moment across markets within the percentile. To be consistent with the data, I simulate 120 labor markets and compute the median of the targeted moment across the different simulations.[11] Table 4 reports the estimated values, the targeted moments (and sources), and the corresponding model moments.

I target the national unemployment rate averaged over the same years as my panel, which is tightly linked to the unemployed offer arrival rate Φ. The on-the-job search intensity ξ targets job-to-job transitions in a low-concentration labor market (defined as a market

[11]This can create discrete jumps in the distance function, as one can see in the graphs.

Name	Param.	Value	Target	Data	Model	Source
Offer arrival rate for unemp. workers	Φ	0.10	Unemp. rate	9.2%	9.2%	INSEE
Search intensity for employed workers	ξ	0.51	Job-to-job transition rate	8.6%	8.4%	BTS
Location of Pareto distribution	μ	0.90	Mean productivity	1.00	1.00	Norm.
Shape of Pareto distribution	k	21.0	Regression of mean wage on concentration	0.050	0.050	BTS
Curvature of vacancy-posting costs	γ	13.9	HHI	0.019	0.019	BTS
Linear vacancy-posting costs	φ_v	30.0	Tightness in the median market	1	1.0	Norm.
Operating cost	c_f	1.10e-03	Nbr. of operating firms in the median market	27	27.0	BTS

Table 4: Estimated parameters and targeted moments.

in the 10th percentile of HHI). Firms draw productivity from a Pareto distribution with lower bound $\underline{z}$ chosen so that the employment-weighted mean productivity equals one, and shape k chosen to match the slope of the regression of the market-level mean wage on the concentration level as reported in Table 1. I use $c(v) = \frac{1}{1+\gamma}(\varphi_v v)^{1+\gamma}$ and estimate γ to match the level of concentration in a low-concentration labor market, and φ_v to normalize the tightness in the median market. Finally, the operating cost c_f targets the number of operating firms in the median labor market (a market in the 50th percentile of HHI).

Most targeted moments are well approximated—including job-to-job flows, despite the model not capturing all transitions. In the data, I count all job-to-job moves, whereas the model only generates moves up the job ladder.[12]

[12]Yet this comes at the expense of a high curvature in the vacancy-posting costs. Appendix F.1 shows that the efficiency properties do not vary much with a lower γ.

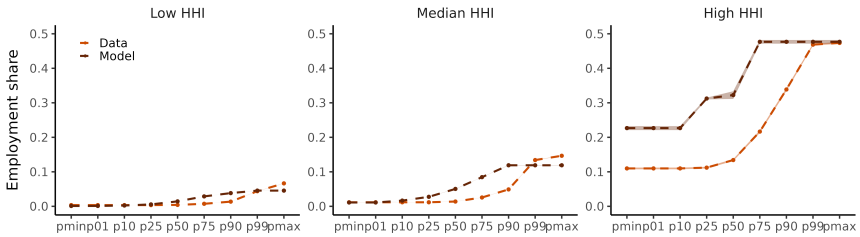


Figure 6: Employment share across the labor share distribution at different concentrations

Variable	Low HHI Data	Low HHI Model	Medium HHI Data	Medium HHI Model	High HHI Data	High HHI Model
Operating cost c_f		0		1.10×10^{-3}		22.00×10^{-3}
HHI^{wb}	0.019	0.019	0.098	0.062	0.369	0.371
Number of firms	154	154	27	27	6	3
Unemployment rate		9.2%		18.9%		41.4%
J2J probability	8.6%	8.4%	7.6%	4.8%	3.0%	1.4%
J2U probability	7.4%	7.4%	6.4%	7.4%	2.5%	7.4%
Total output		0.96		0.94		0.85
Total vacancies		2.9		0.6		0.1
Average wage		0.91		0.92		0.93
Average markdown		0.10		0.13		0.18
Average productivity		1.00		1.06		1.14
$p_{\max}/p_{50}$	20.5	15.8	12.0	4.7	8.0	1.5

Table 5: Non-targeted moments at different concentration levels

Table 5 reports non-targeted moments, and Figure 6 the employment shares across the firms distribution at different concentration levels. Data moments correspond to the median value across markets within a concentration level. Several moments are not observed in the data, in which case the value is omitted in the table.[13]

Although the HHI level is only targeted at low-concentration markets (and the operating cost is chosen to match the number of firms operating in the median market), my model

[13]Unemployment is only available at the 2-digit occupation or commuting zone level.

delivers reasonable levels of concentration and number of firms operating when increasing the operating costs. It is worth noting that, holding all parameters fixed except operating costs, the model cannot match the decline in the J2U probability in more concentrated markets; matching this margin would require allowing job-losing rates to vary across labor markets (e.g., through δ), a margin that is empirically important.[14] In the data, the ratio of employment at the largest firm compared to the median firm tends to be higher than in the model. More generally, my model does not seem to generate enough small firms in concentrated markets. This suggests a thicker productivity distribution (which the model could reconcile with a higher worker's bargaining power). Nevertheless, the model matches the data reasonably well given the restricted set of parameters I estimate.

Before turning to simulations, I assess the sensitivity of the estimation by varying each parameter around its estimated value while holding the others fixed. The distance function displays a clear minimum for each parameter, and the targeted moments respond to movements in their associated parameters. Full sensitivity figures are provided in Appendix E.

4.2 Simulations and wage decomposition

The horse-race between the productivity selection channel and the market power one in the model is underpinned by the shape of the productivity distribution. Figure 7 plots the relationship between concentration and employment-weighted wages for different shapes of the productivity distribution, holding other parameters constant. If the shape of the productivity distribution increases (i.e., the distribution becomes concentrated at its location), the market power dominates, and my model generates a negative slope between concentration and wages, consistent with Berger, Herkenhoff, and Mongey (2022), while the lower the

[14] Bilal (2023) emphasizes that variation in job-losing rates across local labor markets is central to explaining spatial unemployment differences.

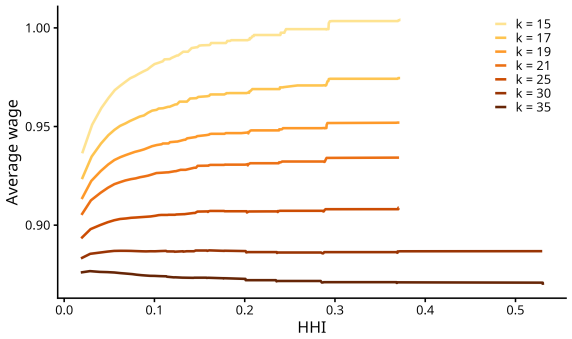


Figure 7: Average wage and concentration in labor markets for different Pareto shapes

shape is (the heavier the tail), the stronger the productivity selection channel becomes as superstar firms drive up workers' outside options. This underlines that, although the operating cost is necessary in generating the positive correlation, it is not sufficient.

I now turn to wage decomposition and markdowns. I quantify how the three terms in the wage equation vary across labor markets with different concentration levels and across firms at different points of the productivity distribution. I simulate 120 labor markets for three values of the operating cost, and report the median value across the different simulations. Within each labor market, I focus on firms at the 0%, 10%, 25%, 50%, 75%, 90%, and 100% productivity percentiles. Table 6 reports, for each firm, total wages, the percentage of the wage accounted for by each term, the implied markdown, and the firm's employment share at three concentration levels (low, median, and high). Appendix E repeats the decomposition for $\theta \in \{0.35, 0.65\}$ (and αs to maintain the Hosios condition) in Tables 14 and 15; results are virtually unchanged.

At high-productivity firms, the worker's share of output accounts for the largest component of wages, while the outside option matters more for workers at low-productivity firms. Despite large variation in firm productivity and wages, markdowns remain surprisingly low, even in concentrated labor markets (between 0.09 at low-productivity firms

Operating cost c_f	0						
Percentile productivity	0%	10%	25%	50%	75%	90%	100%
Firm's productivity	0.90	0.90	0.91	0.93	0.96	1.00	1.17
Total wage	0.81	0.82	0.83	0.85	0.88	0.91	1.01
From output (%)	55.4	55.4	55.2	55.0	54.9	55.1	57.7
From penalty (%)	-7.9	-7.4	-6.7	-5.3	-3.4	-1.8	0.0
From outside option (%)	52.4	52.1	51.5	50.3	48.6	46.7	42.3
Markdown	0.10	0.10	0.09	0.09	0.09	0.09	0.13
Share workers (%)	0.1	0.1	0.2	0.3	0.7	1.7	4.6
HHI				0.019			

Operating cost c_f	1.10×10^{-3}						
Percentile productivity	0%	10%	25%	50%	75%	90%	100%
Firm's productivity	0.98	0.98	0.99	1.01	1.05	1.10	1.17
Total wage	0.85	0.85	0.86	0.88	0.91	0.95	1.00
From output (%)	57.8	57.7	57.5	57.2	57.1	57.6	58.5
From penalty (%)	-6.6	-6.3	-5.5	-4.1	-2.3	-0.9	0.0
From outside option (%)	48.8	48.5	48.0	46.8	45.1	43.4	41.5
Markdown	0.14	0.13	0.13	0.13	0.12	0.13	0.14
Share workers (%)	1.1	1.3	1.6	2.6	5.0	8.5	11.9
HHI				0.062			

Operating cost c_f	22.00×10^{-3}						
Percentile productivity	0%	10%	25%	50%	75%	90%	100%
Firm's productivity	1.10	1.10	1.10	1.12	1.17	1.17	1.17
Total wage	0.90	0.90	0.90	0.92	0.96	0.96	0.96
From output (%)	61.1	61.1	61.1	60.8	60.8	60.8	60.8
From penalty (%)	-3.1	-3.1	-3.1	-1.6	0.0	0.0	0.0
From outside option (%)	41.9	41.9	41.9	40.9	39.2	39.2	39.2
Markdown	0.18	0.18	0.18	0.18	0.18	0.18	0.18
Share workers (%)	22.0	22.0	22.0	31.0	47.0	47.0	47.0
HHI				0.371			

Table 6: Wage decomposition and markdowns by firm's productivity

in non-concentrated markets and 0.18 in concentrated markets). This helps explain why the productivity-selection channel dominates in my model. Nevertheless, consistent with

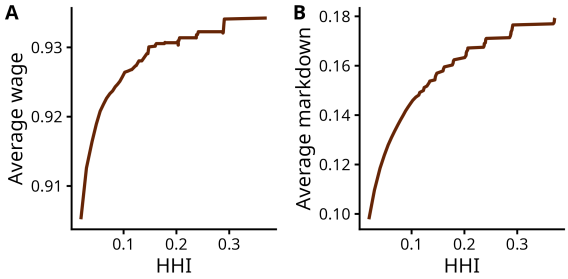


Figure 8: Employment-weighted average wages and markdowns by concentration

intuition, the model delivers one clear pattern: more concentrated markets display both higher average wages and markdowns, as shown in Figure 8, a result also highlighted by Lindenlaub, Oh, and Peters (2024).[15] Notice that markdowns tend to be higher at more productive firms, but this need not be the case due to the poaching penalty which the most productive firms cannot charge.[16]

As a comparison, a few other papers have been measuring markdown levels using different models. Converting their measures to a markdown consistent with my definition of η, Berger, Herkenhoff, and Mongey (2022) reports average markdowns around 0.2 and Deb et al. (2024) around 0.3. My model generates employment-weighted markdown levels between 0.10 and 0.18. This is also smaller than what a back-of-the-envelope comparison to the U.S. labor share (often around 0.6) would suggest.

[15]The figure varies operating costs from 0 to 22×10^{-3}, simulates 120 labor markets, and plots the median employment-weighted average wages (Panel A) and markdowns (Panel B), against the median HHI, across simulations for each operating cost.

[16]Markdowns do not need to be increasing with firm's productivity in the model. The model predicts $w_{i+1} > w_i$ and $z_{i+1} > z_i$ but that does not imply $\eta_{i+1} > \eta_i$. It depends on the productivity gap between i and $i+1$. If $z_i \approx z_{i+1}$, then $\eta_i > \eta_{i+1}$ as i can charge the poaching penalty vis-a-vis $i+1$.

5 Efficiency properties of the model

Previous sections showed that operating costs and vacancy creation shape aggregate output. Higher operating costs screen out small, unproductive firms that crowd the vacancy market, allowing more productive firms to hire more workers, and increasing average productivity. At the same time, a larger number of firms raises total vacancy posting and lowers unemployment, which also increases output. The planner therefore faces a trade-off between the number of firms, the mass of vacancy they post, and the allocation of workers across them.

This section characterizes the first-best allocation and contrasts it with the decentralized equilibrium. It then studies second-best implementations in which entry and production taxes steer firms' operating and vacancy-posting decisions toward the efficient allocation. Overall, these second-best policies come close to the first-best.

5.1 Planner's problem: a first-best solution

A social planner maximizes total output. Relative to the decentralized equilibrium, it can (i) restrict the set of operating firms, (ii) choose vacancy-posting, and (iii) direct worker flows from unemployment and across firms. I characterize the first-best by maximizing the current-value Hamiltonian $\mathcal{H}$. After productivities $\{z_i\}_{i=1}^N$ are realized, the planner chooses the employment allocation $\{n_i\}$, vacancy postings $\{v_i\}$ and v, the aggregate contact rate Λ, and transition rules from unemployment $\mathbb{1}^h$ and between firms $\mathbb{1}^p$ to maximize:

$$\begin{aligned}\mathcal{H} = \max_{\substack{\{n_i\},\{v_i\},v,\Lambda,\\ \mathbb{1}^h,\mathbb{1}^p,\\ N^{\text{firms operating}}}} & \left(\sum_i n_i z_i + \left(1-\sum_i n_i\right) b - \sum_i c(v_i) - c_f N^{\text{firms operating}}\right) \\ & + \sum_i \mu_i \left(-\delta n_i + \left(1-\sum_j n_j\right) \mathbb{1}_i^h \Lambda \frac{v_i}{v} + \sum_k n_k \mathbb{1}_{k\to i}^p \xi \Lambda \frac{v_i}{v} - n_i \xi \Lambda \sum_k \mathbb{1}_{i\to k}^p \frac{v_k}{v} \right)\end{aligned} \tag{6}$$

The Hamiltonian is standard: output plus home production minus vacancy-posting costs, as in Pissarides (2000) and Fukui and Mukoyama (2025). The only additional margin is the total operating costs paid by operating firms, captured by the term $c_f N^{\text{firms operating}}$. The remaining terms impose the laws of motion for employment at each firm, with co-states $\{\mu_i\}$, which are identical to the decentralized environment.

Because the number of operating firms is discrete, the optimal $N^{\text{firms operating}}$ cannot be characterized through a standard first-order condition. Instead, for each candidate N, I solve the planner's problem conditional on having exactly N operating firms. I then choose the N that yields the highest value. Conditional on N, the planner selects the N highest-productivity firms. Appendix D.4 provides more details.

First-order conditions with respect to $\mathbb{1}^h$ and $\mathbb{1}^p$ imply simple transition rules: workers move from unemployment to firm i whenever $\mu_i > 0$, and from i to k whenever $\mu_k > \mu_i$. Taking first-order conditions with respect to $\{n_i\}$, $\{v_i\}$, v, and Λ and rearranging yields an HJB equation for the shadow value of the marginal worker at firm i, together with the planner's optimal vacancy-posting condition:

$$\begin{aligned}(\rho+\delta)\mu_i = z_i - b - 1\cdot\sum_k \mu_k \Lambda \frac{v_k}{v} + \xi\sum_{k>i}\left(1\cdot\mu_k - \mu_i\right)\Lambda\frac{v_k}{v} \\ +(1-\xi)(1-\theta)\sum_k \mu_k \Lambda \frac{\frac{v_k}{v}u + \xi\frac{v_k}{v}\sum_{j<k} n_j - \xi\sum_{j>k}\frac{v_j}{v}n_k}{u+\xi(1-u)}\end{aligned} \tag{7}$$

$$\begin{aligned}c'(v_i) = 1\cdot\frac{\Lambda}{v}\Bigg(\mu_i u + \xi\sum_{j<i}\left(\mu_i-\mu_j\right)n_j \\ -\theta\left(\sum_k \mu_k\left(\frac{v_k}{v}u + \xi\frac{v_k}{v}\sum_{j<k} n_j - \xi\sum_{j>k}\frac{v_j}{v}n_k\right)\right)\Bigg).\end{aligned} \tag{8}$$

These expressions are the planner counterparts to (2) and (1) in the decentralized problem; the differences are highlighted in red.

Equations (2) and (7) differ for two reasons. First, private bargaining does not internalize that a move from unemployment (or to a more productive firm) creates the full social value μ; it only considers the share of the surplus $\alpha\mu$ the worker captures. Second, private agents do not internalize the congestion externality created by matches, captured by the red term in the second line of (7). To build intuition, set $\xi = 0$, so that $(\rho+\delta)\mu_i = z_i - b - \theta \sum_k \mu_k \Lambda \frac{v_k}{v}$. Comparing to the decentralized equation shows that μ_i and S_i coincide if and only if $\theta = \alpha$—the standard Hosios condition (Hosios 1990). With on-the-job search, this congestion term interacts with job-to-job flows.[17]

Equations (1) and (8) also differ along two margins. First, vacancy creation is driven by the private return $(1-\alpha)S_i$, whereas the planner values the full social gain from a match. Second, each firm internalizes the effect of v_i on its own meeting rate, but not the congestion it imposes on other firms. This cross-firm congestion wedge is present even without on-the-job search. Setting $\xi = 0$ highlights the contrast:

$$c'(v_i) = \frac{\Lambda}{v}\left(1 - \theta\frac{v_i}{v}\right) u(1-\alpha)S_i \qquad \text{and} \qquad c'(v_i) = \frac{\Lambda}{v}\left(1 - \theta\sum_k \frac{\mu_k}{\mu_i}\frac{v_k}{v}\right) u\mu_i.$$

In general, efficiency cannot be restored by a simple restriction of the form $\theta = \alpha$, in contrast to the Hosios condition.

More strongly, the first-best allocation cannot, in general, be decentralized even with fully flexible firm-level taxes when there is on-the-job search. The reason is a discrete surplus jump generated by the job ladder: workers climb the job ladder without bilateral wage renegotiation in response to outside offers, and higher-ranked firms earn poaching rents that do not vanish even when productivities are arbitrarily close.

To see this, consider the two most productive firms N and $N-1$ with $z_N = z_{N-1} + \varepsilon$

[17] Fukui and Mukoyama (2025) offers a detailed discussion of the inefficiency created by on-the-job search and how important it is quantitatively.

for some small $\varepsilon > 0$. In the planner's allocation, taking differences in the shadow values (7) and in the vacancy-posting condition (8) yields $\mu_N \approx \mu_{N-1}$ and $v_N \approx v_{N-1}$ as $\varepsilon \to 0$. In the limit, the planner treats firms N and $N-1$ symmetrically. In the decentralized equilibrium, instead, the surplus difference does not vanish:

$$\left(\rho + \delta + \xi\Lambda\frac{v_N}{v}\right)(S_N - S_{N-1}) = (z_N - z_{N-1}) + \xi\Lambda(1-\alpha)S_N\frac{v_N}{v}$$
$$\xrightarrow[\varepsilon\to 0]{} \xi\Lambda(1-\alpha)S_N\frac{v_N}{v} > 0.$$

Because this discrete surplus gap cannot generally be eliminated by firm-level taxes/subsidies when $\xi > 0$, the gap translates into a discontinuity in vacancy-posting incentives. Equivalently, with $\xi > 0$, being ranked above another firm is valuable even when productivities are nearly identical. Under different bargaining protocols (for instance Cahuc, Postel-Vinay, and Robin 2006), this discrete jump could disappear.

To quantify inefficiencies, I simulate 120 markets and solve the decentralized equilibrium and the planner's allocation. Table 7 summarizes the results. The first-best concentrates employment among the most productive firms by sharply reducing the number of operating firms (from 27 to 15). Total vacancies fall, and vacancy-posting becomes more concentrated at the most productive firms. Overall, output increases by 0.8% because production shifts toward more productive firms, even though unemployment is significantly higher.

As a comparison, I solve the decentralized labor market in which firms do not internalize their effect on v. The equations for S_i and J_i are identical, but the vacancy-posting becomes:

$$c'(v_i) = \frac{\Lambda}{v}\left(u + \xi\sum_{j<i} n_j\right)(1-\alpha)S_i,$$

This condition is identical to (1) without the congestion term, so the term in θ disappears.

	Total output	Nbr. firms	HHI^{emp}	Unemp. rate	Total mass of vacancies	Share vacancies by best firm	J2J prob.
Decentralized	0.94	27	0.059	18.9%	0.6	4.1%	4.8%
No congestion internalization	-0.0%	27	0.059	18.9%	0.6	4.1%	4.8%
Planner	+0.8%	15	0.092	24.2%	0.3	7.2%	6.6%

Table 7: Decentralized labor markets and planner's solution

Table 7 shows that firms internalizing congestion barely changes the decentralized equilibrium – total output declines by a negligible amount. Most of the inefficiency comes from the inefficient operation of unproductive firms.

These results remain true when reducing γ to a low value (0.5), as shown in Table 16. The planner sharply reduces the number of operating firms, increasing total output by around 0.6%. The planner concentrates employment at the most productive firms—as expected when the curvature of the vacancy-posting decreases—increasing the HHI relatively more compared to the decentralized market at low gamma. Finally, the fact that firms internalize parts of the congestion effect remains small.

5.2 Entry and production tax

Inefficiency is driven by (i) firm entry and (ii) vacancy-posting decisions. Without on-the-job search ($\xi = 0$), individual taxes and subsidies can restore efficiency: it suffices to choose τ_i so that the decentralized vacancy choices v_i match the planner's.[18] With on-the-job search ($\xi > 0$), firms benefit from poaching and efficiency may require discrete jumps in S that

[18]As the planner's allocation features v_i increasing in productivity, finding $\{\tau_i\}$ that match $\{v_i\}$ in the decentralized economy while preserving the ranking assumption, so that $\{S_i\}$ is increasing in i, is possible. Hence there exists a system $\{\tau_i\}$ that restores efficiency.

	Nbr. firms	HHI^{emp}	Unemp. rate	Total output	Total mass of vacancies	Mean markdown	Mean wage
Decentralized	27	0.059	18.9%	0.94	0.6	0.131	0.92
Entry tax	15	0.092	23.8%	+0.7%	0.3	0.144	0.93
Entry + average tax	15	0.092	24.2%	+0.8%	0.3	0.101	0.78
Planner	15	0.092	24.2%	+0.8%	0.3		

(a) Model outcomes

	Entry tax $(\boldsymbol{\lambda} \cdot c_f)$	Production tax $(\boldsymbol{\tau} \cdot z_i)$
Multipliers	2.22	0.80

(b) Policy multipliers to implement second-best

Table 8: Decentralized labor markets, first-best, and second-best in the median market

cannot generally be replicated with taxes while matching the planner's v_i, as shown above.

Although one cannot generally restore efficiency even with a fully flexible tax system, I investigate how much one achieves by (i) imposing an operating tax/subsidy to incentivize the decentralized equilibrium to select the same operating firms as in the planner's equilibrium, and by (ii) imposing a common output tax/subsidy to ensure the mass of vacancies in both equilibria are the same. This approach does not restore full efficiency, as the composition of vacancy still differs between the planner and the decentralized equilibria.

Table 8 shows that the second-best implementations come close to the planner's output. Imposing an entry tax (by increasing the operating cost by a factor 2.2) selects the same firms in the decentralized equilibrium as in the planner. This increases output by 0.7%. By restricting the number of firms, mean wages increase slightly due to an increase in average productivity. Markdowns increase concomitantly as the outside option of workers decreases. Adding a production tax (decreasing production by 20% to match the mass of vacancies posted in the decentralized market and in the planner) virtually restores efficiency. This

policy decreases the mean wage, as production left to share decreases –although rebating the tax to workers would compensate for the decrease in wages. Table 17 repeats the same exercise in a low- and high-concentrated market, and delivers similar results. This exercise implies that the key policy margin is limiting operation by low-productivity firms.

6 Conclusion

This paper studies the links between labor-market concentration and labor-market outcomes using French administrative data. Contrary to much of the literature, wages rise with labor-market concentration: moving from the 10th to the 90th concentration percentile is associated with a 6.9% increase in daily wages. Worker mobility is lower in concentrated labor markets. These patterns hold in the raw data, in market-level and panel regressions, and are robust across both concentration measures and market definitions.

To interpret these facts, I build on Acemoglu and Hawkins (2014)'s search-and-matching framework and introduce firm granularity and an entry decision: labor markets feature a discrete number of multi-worker firms that post vacancies optimally and endogenously decide whether to operate. The model encompasses two competing forces: market power, and productivity selection, which drive workers' outside options in opposite directions. In the estimation, the productivity selection channel dominates, and mean wages increase with concentration. Markdowns exist but remain small (about 0.10–0.18), and most wage variation reflects workers' share of output.

I then characterize the first-best allocation, in which a planner chooses the set of operating firms, vacancy-posting decisions, and worker allocation across firms. The planner concentrates employment among high-productivity firms, raising output by about 0.8%, at the cost of higher unemployment. Finally, I show that efficiency cannot, in general, be fully

restored. Nevertheless, an increase in operating costs—economically similar to an entry tax—that deters low-productivity firms from operating removes most of the inefficiencies.

That higher productivity can raise wages and concentration at the same time is a simple yet overlooked mechanism in the current literature. This mechanism has important implications for how public policies should address concerns about labor-market concentration. Understanding why concentration is high in a given labor market, rather than treating concentration itself as the problem, should therefore be a first-order priority. The framework developed here can be used to quantify which concentration changes reflect efficient selection versus distortions that call for policy intervention.

References

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

Acemoglu, Daron. 2001. "Good jobs versus bad jobs." *Journal of labor Economics* 19 (1): 1–21.

Acemoglu, Daron, and William B Hawkins. 2014. "Search with multi-worker firms." *Theoretical Economics* 9 (3): 583–628.

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. 2022. "Income and wealth distribution in macroeconomics: A continuous-time approach." *The Review of Economic Studies* 89 (1): 45–86.

Autor, David, Christina Patterson, and John Van Reenen. 2023. *Local and national concentration trends in jobs and sales: The role of structural transformation.* Technical report. National Bureau of Economic Research.

Azar, José, Ioana Marinescu, and Marshall Steinbaum. 2022. "Labor market concentration." *Journal of Human Resources* 57 (S): S167–S199.

Azar, José, Ioana Marinescu, Marshall Steinbaum, and Bledi Taska. 2020. "Concentration in US labor markets: Evidence from online vacancy data." *Labour Economics* 66:101886.

Bassanini, Andrea, Cyprien Batut, and Eve Caroli. 2023. "Labor market concentration and wages: Incumbents versus new hires." *Labour Economics* 81:102338.

Bassanini, Andrea, Giulia Bovini, Eve Caroli, Jorge Casanova Ferrando, Federico Cingano, Paolo Falco, Florentino Felgueroso, Marcel Jansen, Pedro S Martins, António Melo, et al. 2024. "Labor market concentration, wages and job security in Europe." *Journal of Human Resources.*

Bergé, Laurent. 2018. "Efficient estimation of maximum likelihood models with multiple fixed-effects: the R package FENmlm." *CREA Discussion Papers,* no. 13.

Berger, David, Kyle Herkenhoff, Andreas R Kostøl, and Simon Mongey. 2024. "An anatomy of monopsony: Search frictions, amenities, and bargaining in concentrated markets." *NBER Macroeconomics Annual* 38 (1): 1–47.

Berger, David, Kyle Herkenhoff, and Simon Mongey. 2022. "Labor market power." *American Economic Review* 112 (4): 1147–1193.

Bilal, Adrien. 2023. "The geography of unemployment." *The Quarterly Journal of Economics* 138 (3): 1507–1576.

Boar, Corina, and Virgiliu Midrigan. 2025. "Markups and inequality." *Review of Economic Studies* 92 (5): 2828–2860.

Burdett, Kenneth, and Dale T Mortensen. 1998. "Wage differentials, employer size, and unemployment." *International Economic Review,* 257–273.

Burstein, Ariel T, Vasco M Carvalho, and Basile Grassi. 2025. "Bottom-up markup fluctuations." *The Quarterly Journal of Economics* 140 (4): 2619–2684.

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.

Cestone, Giacinta, Chiara Fumagalli, Francis Kramarz, and Giovanni Pica. 2024. "Exploiting growth opportunities: The role of internal labour markets." *Review of Economic Studies* 91 (5): 2676–2716.

Correia, Sergio. 2016. "A feasible estimator for linear models with multi-way fixed effects."

Deb, Joyee, Jan Eeckhout, Neel Patel, and Jesse M Shapiro. 2024. "Market Power and Wage Inequality." *Econometrica* 92 (3): 603–636.

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu. 2015. "Competition, markups, and the gains from international trade." *American Economic Review* 105 (10): 3183–3221.

Fukui, Masao, and Toshihiko Mukoyama. 2025. *Efficiency in Job-Ladder Models.* NBER Working Paper 33910. National Bureau of Economic Research, June.

Hopenhayn, Hugo A. 1992. "Entry, exit, and firm dynamics in long run equilibrium." *Econometrica,* 1127–1150.

Hosios, Arthur J. 1990. "On the efficiency of matching and related models of search and unemployment." *The Review of Economic Studies* 57 (2): 279–298.

Jarosch, Gregor, Jan Sebastian Nimczik, and Isaac Sorkin. 2024. "Granular search, market structure, and wages." *Review of Economic Studies* 91 (6): 3569–3607.

Lindenlaub, Ilse, Ryungha Oh, and Michael Peters. 2024. "Spatial Firm Sorting and Local Monopsony Power." Working paper.

Manning, Alan. 2003. “The real thin theory: monopsony in modern labour markets.” *Labour Economics* 10 (2): 105–131.

Marinescu, Ioana, Ivan Ouss, and Louis-Daniel Pape. 2021. “Wages, hires, and labor market concentration.” *Journal of Economic Behavior & Organization* 184:506–605.

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

Menzio, Guido, and Shouyong Shi. 2010. “Block Recursive Equilibria for Stochastic Models of Search on the Job.” *Journal of Economic Theory* 145, no. 4 (July): 1453–1494.

Mortensen, Dale T, and Christopher A Pissarides. 1994. “Job creation and job destruction in the theory of unemployment.” *The Review of Economic Studies* 61 (3): 397–415.

Nimczik, Jan Sebastian. 2020. “Job mobility networks and data-driven labor markets.” Working paper.

Palladino, Marco Guido, Damien Babet, and Olivier Godechot. 2025. “In the Land of AKM: Explaining the Dynamics of Wage Inequality in France.”

Pissarides, Christopher A. 2000. *Equilibrium unemployment theory*. MIT Press.

Postel–Vinay, Fabien, and Jean–Marc Robin. 2002. “Equilibrium wage dispersion with worker and employer heterogeneity.” *Econometrica* 70 (6): 2295–2350.

Rinz, Kevin. 2022. “Labor market concentration, earnings, and inequality.” *Journal of Human Resources* 57 (S): S251–S283.

Schmutte, Ian M. 2014. “Free to move? A network analytic approach for learning the limits to job mobility.” *Labour Economics* 29:49–61.

Stole, Lars A., and Jeffrey Zwiebel. 1996. “Intra-firm bargaining under non-binding contracts.” *The Review of Economic Studies* 63 (3): 375–410.

Sun, Liyang, and Sarah Abraham. 2021. “Estimating dynamic treatment effects in event studies with heterogeneous treatment effects.” *Journal of Econometrics* 225 (2): 175–199.

Yeh, Chen, Claudia Macaluso, and Brad Hershbein. 2022. “Monopsony in the US labor market.” *American Economic Review* 112 (7): 2099–2138.

A Raw data appendix

A.1 Data construction and caveats

BTS and PTS are constructed from compulsory employer declarations used for social contributions, which gives employers strong incentives to report accurately. Worker identifiers in BTS are reshuffled between waves, which prevents following individuals outside a given $(t-1, t)$ window. [19]

BTS is a repeated cross-section of all jobs at the establishment level. Each wave t contains information for years $(t-1, t)$. For each year, I use the year-t part of wave t to construct labor-market averages (concentration, wages, demographics, etc.). I use the change between $(t-1)$ and t within the same wave to compute transition probabilities for year $(t-1)$. Occupations were not systematically recorded prior to 2009 in BTS, which motivates restricting the main analysis to 2009–2019.

For the 2016 wave (covering 2015–2016), roughly 10% of variables are missing. Based on communication with INSEE, this is likely due to a change in the way data were reported to them. As a result, I use the 2016 wave to compute transition probabilities for 2015 (as usual), but I compute labor-market averages for 2016 using the year-$(t-1)$ information from the 2017 wave (i.e., 2016 measured in the 2017 wave).

I restrict the analysis to private-sector workers, keep full-time, regular jobs (excluding interns and state-subsidized jobs), restrict to workers aged 20–60, and drop jobs paying less than one monthly minimum wage over the year.[20] These restrictions leave a large sample of

[19] Methods for probabilistic matching across waves have been developed Palladino, Babet, and Godechot (2025) but are not used here.

[20] I exclude civil servants and civil-service occupations (such as firefighters, police officers, or military personnel) using employer-category and civil-servant-status variables. Using 25 as the lower age cutoff yields similar results. I drop unemployment spells recorded in the data; they correspond to benefit claims and therefore do not capture all unemployment

13–16 million worker-year observations annually out of around 18–19 million employed in the private sector. Many of the excluded observations correspond to workers only marginally attached to the labor force.

To compute concentration at the firm level, I merge legal units into firms using CEP.[21] This merge is imperfect early in the sample, when CEP was still being rolled out and initially focused on profiling the largest firms (so that legal units were more likely to be merged for the largest firms). The coverage stabilizes later on and becomes consistent. Using CEP typically increases measured concentration, but it does not seem to materially affect the results relative to computing concentration at the legal-unit level.

To compute unemployment and transitions, I count as non-employment all days not covered by a job spell between two consecutive main spells. I record an unemployment spell if the uncovered days sum to at least 60. I classify workers as stayers if they keep the same main employer with no unemployment spell; as job-to-job (J2J) movers at the end of t if they change main employer with no unemployment spell between t and $t+1$; as job loss (J2U) in t if they change main employer in $t+1$ after an unemployment spell; and as job finding (U2J) in $t+1$ if they start a new main spell after an unemployment spell.[22] Results are robust to alternative thresholds.

A final note on occupation choice: although not an official category, 3-digit occupations have a sensible meaning. For example, within blue-collar workers (1-digit, PCS 6), drivers (2-digit, PCS 64) bunch together truck and bus drivers (4-digit, PCS 641a and 641b) as PCS

spells.

[21] See Cestone et al. (2024) for evidence that internal labor markets within business groups differ from external labor markets.

[22] For example, a worker with a main job, then a 70-day job, and then a new main job is classified as J2J (no uncovered time). In contrast, a worker with a main job, then 35 days without work, then a short job, then 35 days without work, and then a new main job is classified as J2U and U2J (70 uncovered days).

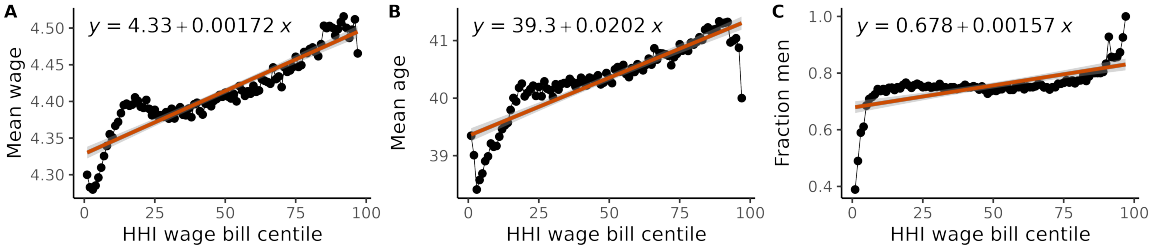


Figure 9: Median of labor market averages by HHI percentile

641 and taxi and other car drivers (PCS 642a and 642b) as PCS 642.

A.2 Raw relationships between concentration, wages, and worker composition

Using the raw BTS data, I compute labor-market averages (wages and demographics) and transition probabilities as described in Section 1. I then group labor markets into HHI percentiles and plot, for each percentile, the median of these market-level moments.

Panel A of Figure 9 (the raw wage–concentration relationship) is discussed in the main text. Panels B–C show that more concentrated labor markets also differ systematically in worker composition: average age is higher (Panel B) and the fraction of men is higher (Panel C).[23] These sorting patterns motivate the worker- and job-level controls used in the regressions.

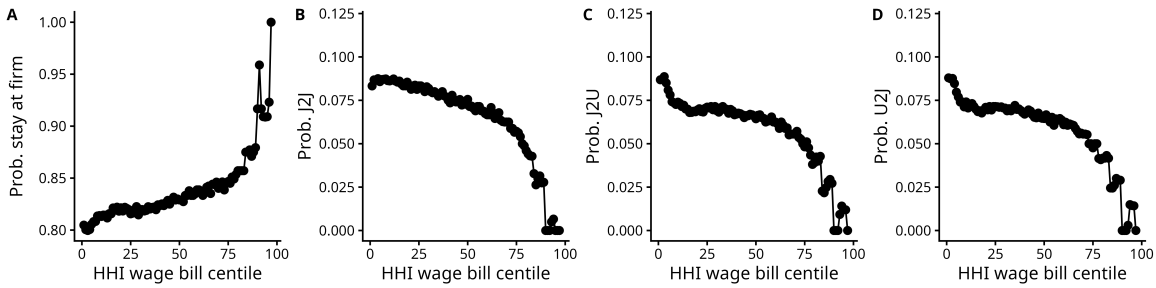


Figure 10: Median transition probabilities by HHI percentile

A.3 Raw relationships between concentration and worker flows

Panel B of Figure 10 (job-to-job transitions) is discussed in the main text. As expected, the probability of staying with the same main employer (Panel A) and the probability of moving from unemployment to a job (U2J, Panel D) both decline with concentration, consistent with fewer firms operating in more concentrated markets. The probability of a job-to-unemployment transition (J2U, Panel C) also declines with concentration, although this pattern is less straightforward to interpret.

B Additional regression specifications

B.1 Market-level regressions

Table 9 extends Table 1 by using HHI^{emp} and different fixed-effect. Columns (i)–(ii) omit the additional market-level controls, while Columns (iii)–(iv) include them. The first panel uses occupation $\times$ commuting-zone fixed effects (baseline), while the second uses occupation $\times$

[23]The fraction of men might seem extremely high, yet: (i) I exclude civil servants, who disproportionately employ women; and (ii) non-concentrated markets tend to be larger and to employ more women. Combining these two effects explains the high fraction of men.

	HHI^{wb} (i)	HHI^{emp} (ii)	HHI^{wb} (iii)	HHI^{emp} (iv)
Occupation × commuting-zone fixed effects				
Concentration	0.0820***	0.0436***	0.0500***	0.0066
	(0.0058)	(0.0058)	(0.0060)	(0.0065)
Occ.×CZ FEs	✓	✓	✓	✓
Add. controls			✓	✓
Observations	407K	407K	407K	407K
R^2	0.905	0.904	0.909	0.908
Within R^2	0.006	0.002	0.045	0.043
Occupation × year fixed effects				
Concentration	0.0390***	-0.0100**	0.0084	-0.0530***
	(0.0049)	(0.0050)	(0.0052)	(0.0054)
Occ.×year FEs	✓	✓	✓	✓
Add. controls			✓	✓
Observations	407K	407K	407K	407K
R^2	0.857	0.856	0.864	0.865
Within R^2	0.002	0.000	0.056	0.058

Clustered (LM) standard errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 9: Market-level regressions: mean log daily wage on HHI measures

year fixed effects. In the baseline specification with occupation × commuting-zone fixed effects, wages are positively associated with both HHI^{wb} and HHI^{emp}. Under occupation × year fixed effects, the coefficient on employment-share concentration becomes negative—the only regressions in which I obtain a negative correlation between wages and concentration.

	Log daily wage (vi)	Stay at firm (vii)	J2J (viii)	J2U (ix)	U2J (x)
HHI^{wb}	0.1312***	0.0194***	-0.0286***	-0.0184***	-0.0058***
	(0.0120)	(0.0060)	(0.0086)	(0.0047)	(0.0049)
Sex & Occ.×CZ FEs	✓	✓	✓	✓	✓
Add. controls	✓	✓	✓	✓	✓
Observations	8.7M	7.2M	7.2M	7.2M	7.2M
R^2	0.568	0.121	0.028	0.105	0.085
Within R^2	0.144	0.067	0.014	0.054	0.047

Clustered (firm) standard errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 10: Panel data regressions: log daily wage and transitions with tenure and work-experience controls

B.2 Panel data regressions

Table 10 complements Table 2 in the main text[24] by adding worker-level controls for tenure and work experience (years worked). These controls absorb lifecycle and match-quality differences that may covary with concentration; the estimates remain close to the baseline panel specification.

Table 11 extend Tables 2 and 10 by varying the concentration measure (HHI^{wb} versus HHI^{emp}) and the set of controls; columns (i)–(iv) focuses on wages, while columns (v)–(viii) focuses on job-to-job transitions. Columns (i)–(ii) and (v)–(vi) report the baseline controls, while Columns (iii)–(iv) and (vii)–(viii) add tenure and work-experience controls. In each table, the top panel reports specifications with sex and occupation $\times$ commuting-zone fixed effects, and the bottom panel replaces these with worker and firm fixed effects.

[24]All panel data regressions are performed using the R package fixest developed by Bergé (2018). The algorithm uses alternate projection (as in Correia (2016)) to remove fixed effects.

	Daily wage				J2J probability			
	HHI^{wb}	HHI^{emp}	HHI^{wb}	HHI^{emp}	HHI^{wb}	HHI^{emp}	HHI^{wb}	HHI^{emp}
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Sex and occupation × commuting-zone fixed effects								
Concentration	0.1380***	0.0189*	0.1312***	0.0203*	-0.0314***	-0.0203**	-0.0286***	-0.0196**
	(0.0120)	(0.0106)	(0.0120)	(0.0113)	(0.0082)	(0.0086)	(0.0086)	(0.0093)
Sex & Occ.×CZ FEs	✓	✓	✓	✓	✓	✓	✓	✓
Add. controls			✓	✓			✓	✓
Observations	8.7M	8.7M	8.7M	8.7M	7.2M	7.2M	7.2M	7.2M
R^2	0.560	0.560	0.568	0.568	0.025	0.025	0.028	0.028
Within R^2	0.127	0.127	0.144	0.144	0.011	0.011	0.014	0.014
Worker and firm fixed effects								
Concentration	0.0259***	0.0087	0.0266***	0.0096	-0.0146***	-0.0124***	-0.0129***	-0.0105***
	(0.0065)	(0.0072)	(0.0066)	(0.0072)	(0.0035)	(0.0037)	(0.0036)	(0.0038)
Worker FEs	✓	✓	✓	✓	✓	✓	✓	✓
Firm FEs	✓	✓	✓	✓	✓	✓	✓	✓
Add. controls			✓	✓			✓	✓
Observations	8.7M	8.7M	8.7M	8.7M	7.2M	7.2M	7.2M	7.2M
R^2	0.900	0.900	0.900	0.900	0.406	0.406	0.409	0.409
Within R^2	0.038	0.038	0.038	0.038	0.002	0.002	0.007	0.007

Clustered (firm) standard errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 11: Panel data regressions on HHI measures: daily wage and J2J probability

Across these panel specifications, higher concentration is associated with higher wages and lower job-to-job mobility; results are similar when using employment-share HHI instead of wage-bill HHI, and when adding tenure and work-experience controls. When using worker and firm fixed effects, magnitudes of coefficients on wages drop significantly (consistent with firm's productivity representing a large share of the wage, as is discussed in the model), but remain consistent with the main results.

C Event-study appendix

C.1 Event-study design

Building on the event-study setup described in Subsection 2.3, I restrict attention to labor markets with at least 300 workers in all years, to ensure that small markets do not mechanically drive measured changes in concentration. I define a shock as an increase (or decrease) in HHI^{wb} greater than 0.1 over two consecutive years, and I drop markets with multiple shocks. Control markets are defined within the same occupation, in commuting zones that are not geographically adjacent to treated markets, and with consecutive changes in HHI^{wb} below 0.05 throughout the years. This procedure yields 159 markets with a single large increase and 108 with a single large decrease, along with about 6,000 control markets (out of roughly 34,000 occupation $\times$ commuting-zone cells).

Treated events are spread across years, commuting zones, and occupations, and are not driven by a small set of locations, occupations or a single period. Because I restrict attention to relatively large labor markets (at least 300 workers in all years), treated markets are mechanically more prevalent in larger, relatively common occupations (e.g., manual/blue-collar occupations). Event-study regressions include commuting-zone and occupation-by-year fixed effects to further alleviate concerns about cross-sectional differences across

locations and time.

C.2 Dominant-firm decomposition following HHI increases

Figure 11 decomposes treated labor markets following large HHI increases into (i) the firm that is largest after the shock (the "dominant" firm) and (ii) all other firms in the same market. Panel A shows that the dominant firm expands sharply at the time of the shock (by roughly 250 workers relative to year -1). Panel B plots average wages at the dominant firm, at the remaining firms in the treated market, and in control markets. Wages are higher at the dominant firm even before the shock and, if anything, increase further relative to control markets after the shock; wages at other firms in the treated market also rise modestly. Overall, this evidence suggests that large HHI increases are primarily driven by the rapid expansion of a high-wage firm, and tend to increase wages within the shocked market.

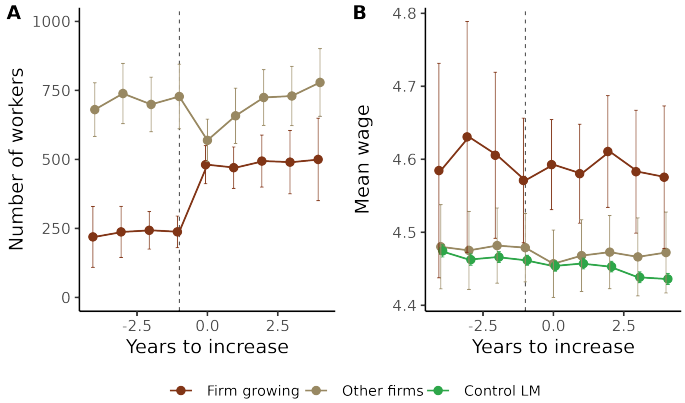


Figure 11: Growing firm versus the rest of the labor market following the HHI increase

C.3 Worker-level matched event-study around HHI increases

To assess how these market-level changes map into individual outcomes, I repeat the exercise at the worker-level. I classify workers based on their firm and location in year -1 (the year before the shock) into three groups: (i) workers employed at the firm that becomes dominant after the shock; (ii) workers in the same treated labor market but employed at other firms (excluding firms in the dominant firm's business group); and (iii) workers in labor markets adjacent to the treated market (again excluding the dominant firm's business group). I match each group to control workers based on sex and quintiles of tenure, age, work experience, and firm size. To isolate changes in workers' outside options within a labor market, I keep workers only for as long as they remain in their year -1 labor market. The resulting sample includes 574 workers at dominant firms, about 6,000 workers at other firms in treated markets, 47,000 workers in adjacent markets, and 350,000 matched control workers.

I estimate event-study regressions on this matched worker sample, focusing on log daily wages and the probability of a job-to-job transition in the next period. For wages, I normalize event time at year -1 (the last pre-shock year); for job-to-job transitions, I normalize at year

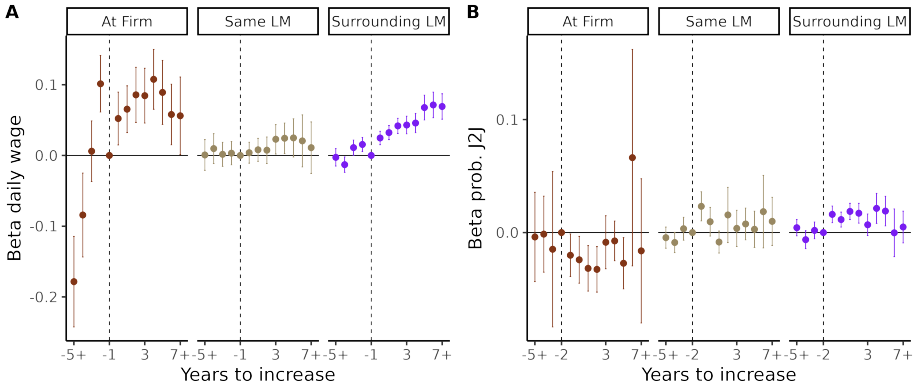


Figure 12: Large increases in concentration over workers' wages and J2J probability

-2. Because the analysis conditions on a large within-market change between two years, I do not include labor-market fixed effects. Following Berger et al. (2024), I instead include occupation-by-year fixed effects, so identification comes from comparing treated and control commuting zones within the same occupation-year. I also include commuting-zone fixed effects, as well as a sex indicator and an age polynomial. The regressions are:

$$\text{daily wage}_{i,t} = \sum_{k \neq -1} \beta_k^{\text{wage}} \cdot \mathbb{1}\{t - \tau_i = k\} + \mu_{o(i)t} + \gamma_{c(i)} + controls + \epsilon_{it}$$

$$\text{J2J trans}_{i,t} = \sum_{k \neq -2} \beta_k^{\text{J2J}} \cdot \mathbb{1}\{t - \tau_i = k\} + \mu_{o(i)t} + \gamma_{c(i)} + controls + \epsilon_{it}$$

Because spillovers across commuting zones (e.g., worker mobility) can contaminate treated and not-yet-treated units, I estimate the event-study using the Sun and Abraham (2021) correction for staggered adoption. Standard errors are clustered at the firm level; results are similar when using standard two-way fixed-effects event-study specifications.

Figure 12 reports worker-level event-study coefficients for log daily wages (Panel A) and the probability of a job-to-job transition (Panel B), separately for the three worker groups. For workers at the dominant firm, Panel A exhibits a pronounced positive pre-trend,

consistent with the firm improving and raising wages prior to the large expansion that generates the HHI shock; for this reason, post-shock wage gains for this group should be interpreted with caution. Nonetheless, wages at the dominant firm increase further after the shock (roughly 5%–10% relative to year -1), while job-to-job mobility declines in the years following the shock. For workers at other firms in the treated market, wage effects are smaller (around 2% after a few years), and job-to-job transitions increase on impact. Workers in adjacent markets display rising wages and a modest increase in job-to-job transitions over time.

These worker-level patterns are consistent with the dominant firm poaching workers from other firms in the same market, as suggested by the decline in employment at non-dominant firms in Figure 11. In addition, wage increases at other firms are not solely driven by workers switching to the dominant firm: results are similar when restricting the sample to workers who remain at the same firm and in the same labor market as in year -1, which isolates a within-market outside-option channel.

Taken together, the market- and worker-level evidence points to large HHI increases being driven by rapid expansion of a high-wage firm. Concentration rises while wages increase and job-to-job mobility declines.

D Model appendix

D.1 Surplus and wage derivations

I derive the surplus and wage equations from the HJB equations and the bargaining protocol, following Acemoglu and Hawkins (2014).

The marginal surplus is $S_i(n_i) \equiv \frac{\partial J_i(n_i)}{\partial n} + V_i(n_i) - U$. Bargaining implies $V_i(n_i) - U = \alpha S_i(n_i)$ and $\frac{\partial J_i(n_i)}{\partial n} = (1-\alpha)S_i(n_i)$, so $\mathbb{1}_{\{V_i > U\}} = \mathbb{1}_{\{S_i > 0\}}$ and $\mathbb{1}_{\{V_j > V_i\}} = \mathbb{1}_{\{S_j > S_i\}} = \mathbb{1}_{j>i}$.

start by taking the difference between $V_i(n_i)$ and U. use bargaining to replace $V_i - U$ by $\frac{\alpha}{1-\alpha}\frac{\partial J_i(n_i)}{\partial n}$,[25] apply the transition rules, and use $\lambda^j = \xi\Lambda\frac{v_j}{v}$ to obtain:

$$\begin{aligned}(\rho+\delta)\frac{\alpha}{1-\alpha}\frac{\partial J_i(n_i)}{\partial n} &= w_i(n_i) + \frac{\alpha}{1-\alpha}\xi\Lambda\sum_{j>i}^{N}\left(\frac{\partial J_j(n_j)}{\partial n} - \frac{\partial J_i(n_i)}{\partial n}\right)\frac{v_j}{v} - \rho U \\ &+ \left(\Big(-\xi\Lambda\sum_{j>i}^{N}\frac{v_j}{v} - \delta\Big)n_i + \Lambda\frac{v_i}{v}\Big(u + \xi\sum_{j<i}^{N} n_j\Big)\right)\frac{\alpha}{1-\alpha}\frac{\partial^2 J_i(n_i)}{\partial n^2}.\end{aligned} \tag{9}$$

Taking the derivative of the firm's HJB equation with respect to n (using the envelope theorem) yields:

$$\begin{aligned}\rho\frac{\partial J_i(n_i)}{\partial n} &= z_i - w_i(n_i) - \frac{\partial w_i(n_i)}{\partial n}n_i + \left(-\xi\Lambda\sum_{j>i}^{N}\frac{v_j}{v} - \delta\right)\frac{\partial J_i(n_i)}{\partial n} \\ &+ \left(\Big(-\xi\Lambda\sum_{j>i}^{N}\frac{v_j}{v} - \delta\Big)n_i + \Lambda\frac{v_i}{v}\Big(u + \xi\sum_{j<i}^{N} n_j\Big)\right)\frac{\partial^2 J_i(n_i)}{\partial n^2}.\end{aligned} \tag{10}$$

multiply (10) by $\frac{\alpha}{1-\alpha}$, compare with (9), and rearrange to obtain:

$$w_i + \alpha\frac{\partial w_i(n_i)}{\partial n}n_i = \alpha z_i + (1-\alpha)\rho U - \alpha\Lambda\xi\sum_{j>i}^{N}\frac{\partial J_j(n_j)}{\partial n}\frac{v_j}{v},$$

Assuming U and the aggregate poaching term $\sum_{j>i}^{N}\frac{\partial J_j(n_j)}{\partial n}\frac{v_j}{v}$ are independent of n_i, the right-hand side is constant in n_i and the wage ODE is degenerate, so w_i is independent of n_i. Constant returns to scale in production are crucial for this result. Using the unemployed

[25] I assume that $\frac{\partial U}{\partial n} = 0$ as mentioned in the main text, so that $\frac{\partial V_i(n_i)}{\partial n} = \frac{\alpha}{1-\alpha}\frac{\partial^2 J_i(n_i)}{\partial n^2}$.

worker's HJB and bargaining then gives:

$$w_i = \alpha z_i + (1-\alpha)b + (1-\alpha)\alpha\Lambda \sum_{j=1}^{N} S_j \mathbb{1}_{\{S_j>0\}} \frac{v_j}{v} - (1-\alpha)\alpha\Lambda\xi \sum_{j>i}^{N} S_j \frac{v_j}{v}.$$

which is (3) from the main text.

As w_i is independent of n_i, (10) becomes:

$$\rho \frac{\partial J_i(n_i)}{\partial n} = z_i - w_i + \left(-\xi\Lambda \sum_{j>i}^{N} \frac{v_j}{v} - \delta \right) \frac{\partial J_i(n_i)}{\partial n} + \dot{n}_i \frac{\partial^2 J_i(n_i)}{\partial n^2}. \tag{11}$$

Guessing that $\frac{\partial J_i(n_i)}{\partial n}$ is constant in n_i solves (11) (again treating $\sum_{j>i}^{N} \frac{v_j}{v}$ as independent of n_i), so the drift term drops. By the same logic, V_i and therefore S_i are independent of n_i. Summing $V_i - U$ and (11) yields:

$$(\rho+\delta)S_i = z_i - \xi\Lambda \sum_{j>i}^{N} \frac{v_j}{v} \frac{\partial J_i(n_i)}{\partial n} + \xi\Lambda \sum_{j>i}^{N} (V_j - V_i) \frac{v_j}{v} - \rho U$$

I finally use the unemployed worker's HJB to substitute for ρU, and apply the transition rules and bargaining protocol to obtain (2). Since S_i is increasing in z_i, an equilibrium with a productivity-ordered job ladder exists, as assumed in the main text. I cannot, however, prove it is the unique equilibrium. There is no *a priori* reason why one could not find equilibria in which workers are better off in less productive firms.

D.2 Event-study exercise in the model

I study how a productivity increase at one firm affects its labor market in a setting that mirrors the event-study in the data. For simplicity, I abstract from endogenous entry and

	At most productive firm	At other firms
Wage change	8.02%	0.12%
J2J prob. change	0.00%	0.78%

Table 12: Event-study in the model

exit—I assume that firms operating before the shock remain active after[26]—and I assume that the productivity ranking is unchanged so that poaching decisions are unaffected.

Because the model is stationary, I cannot compute transition dynamics. Instead, I compare two steady states. I start from a draw of productivities $\{z_i\}_{i=1}^N$ (restricting to the operating firms) and the associated equilibrium objects $(\{S_i\}_{i=1}^N, \{v_i\}_{i=1}^N, \Lambda, \{n_i\}_{i=1}^N, u, v, \{w_i\}_{i=1}^N)$. I then introduce a small change $\mathrm{d}z_i$ and take first-order perturbations of (1), (2), and (3), together with the definitions of Λ and v and the stationary laws of motion for $\{n_i\}$ and u, to obtain a system of linear equations in $(\{\mathrm{d}S_i\}_{i=1}^N, \{\mathrm{d}v_i\}_{i=1}^N, \mathrm{d}\Lambda, \{\mathrm{d}n_i\}_{i=1}^N, \mathrm{d}u, \mathrm{d}v, \{\mathrm{d}w_i\}_{i=1}^N)$ (and $\{\mathrm{d}\eta_i\}_{i=1}^N$ and $\{\mathrm{dJ2J}_i\}_{i=1}^N$)—omitted for brevity.

I solve the system by matrix inversion and feed a 13.2% increase in the dominant firm's productivity (chosen to match an 8% increase in wages at the dominant firm) and no productivity change at other firms. As reported in Table 12, average (employment-weighted) wages at other firms increase by around 0.1% through the higher outside option. Vacancy posting and worker flows reallocate toward the dominant firm: there is no change in J2J probability at the highest firm because workers climb the job ladder, but other workers are 0.8% more likely to move up the ladder. These patterns are qualitatively consistent with the event-study in the data (Appendix C), although smaller in magnitude.

The intuition is as follows. The dominant firm becomes more profitable and posts

[26]In the data event-study, the number of firms decreases slightly about a year after the shock. In the model, a productivity increase at one firm reduces other firms' match surplus by raising workers' outside option, potentially making marginal firms unprofitable.

more vacancies, increasing inflows from unemployment and poaching from other firms. At the same time, the dominant firm's higher surplus raises workers' outside option, which puts upward pressure on wages at all firms; however, in the current calibration, these outside-option wage spillovers to non-shocked firms are smaller than in the data.

The entry of a new firm would have analogous effects on wages and transitions by increasing workers' outside option, and by reducing the share of vacancies each firm is posting.

D.3 Sketch of numerical solution and estimation

I present here a sketch of the algorithm used to solve for the equilibrium with operating costs. Because each firm's surplus and vacancy posting depend on all other firms, I solve for the equilibrium each time the set of operating firms changes. I draw $N = 154$ productivities (to match the number of firms operating in a low-concentration market) and look for a set of operating firms such that all operating firms make non-negative profits. I proceed by successive iterations: I first solve the labor market assuming all firms operate. I then compute each firm's stationary value for J_i^* given the other firms' productivities and the operating cost. If some firm cannot cover the operating cost, I drop the lowest-productivity firm, recompute the equilibrium, and repeat until all remaining firms operate.

For each numerical exercise in the paper, I simulate 120 independent labor markets (independent productivity draws) and solve each market separately. I fix the random seed so that the same 120 productivity draws are used across the estimation and all numerical simulations. Productivity draws are never tied due to random draws from a continuous distribution, so the job ladder is always well defined. Model moments are computed as medians across simulated markets (consistent with the data), and I estimate parameters by

minimizing

$$\sum_i \frac{(\text{median model moment}_i - \text{targeted moment}_i)^2}{\text{targeted moment}_i^2}.$$

The numerical algorithm is:

1. Draw productivities and rank firms from the lowest (1) to the highest (N);
2. Initialize with all firms operating;
3. Guess $\{v_i\}$, $\{S_i\}$ and the worker distribution, and compute the implied contact rates;
4. Given contact rates, solve for $\{S_i\}$ using a finite-difference scheme as in Achdou et al. (2022);
5. Given $\{S_i\}$ and the worker distribution, update $\{v_i\}$ by iterating firms' best responses (each firm solves for its optimal v_i given v_{-i}) and recompute contact rates;
6. Given contact rates and $\{S_i\}$, update the stationary worker distribution (and u);
7. Iterate steps 4–6 until convergence;
8. Recover wages and compute $\{J_i^*\}$. If the lowest-productivity operating firm has $J_i^* < 0$, drop it and return to step 3; otherwise stop.

D.4 Planner's problem

I show how to obtain the system of equations characterizing the planner's problem (7)–(8) from the Hamiltonian (6). Because the number of operating firms is discrete, the optimal $N^{\text{firms operating}}$ cannot be characterized through a standard first-order condition. Instead, for each candidate N, I solve the planner's problem conditional on having exactly N operating firms. From the FOCs with respect to $\mathbb{1}_i^h$ and $\mathbb{1}_{i\to k}^p$, matches from unemployment are created

only if $\mu_i > 0$, and poaching from i to k occurs only if $\mu_k > \mu_i$. Replace v by $\sum_i v_i$ and $\frac{\Lambda}{v}$ by $\Phi^\theta \frac{\left(\sum_i v_i\right)^{-\theta}}{\left(1-\sum_i n_i+\xi\sum_i n_i\right)^{1-\theta}}$ in $\mathcal{H}^N$. Then take the FOC with respect to v_i to obtain:

$$c'(v_i) = \mu_i \frac{\Lambda}{v} u \mathbb{1}_i^h + \mu_i \left(\sum_k n_k \mathbb{1}_{k\to i}^p\right) \xi \frac{\Lambda}{v} - \xi \left(\sum_k \mu_k n_k \mathbb{1}_{k\to i}^p\right) \frac{\Lambda}{v}$$
$$- \theta \sum_k \mu_k u \mathbb{1}_k^h \frac{\Lambda}{v}\frac{v_k}{v} - \theta \sum_k \xi \mu_k \left(\sum_j n_j \mathbb{1}_{j\to k}^p\right) \frac{\Lambda}{v}\frac{v_k}{v} + \theta \sum_k \xi \mu_k n_k \left(\sum_j \mathbb{1}_{k\to j}^p \frac{\Lambda}{v}\frac{v_j}{v}\right)$$

where the first line captures the direct effect of v_i on contacts and vacancy-posting costs, and the second line captures the indirect congestion effect through Λ/v. Rearranging yields (8).

Taking the FOC with respect to n_i yields:

$$\begin{aligned}\rho\mu_i = {} & z_i - b - \delta\mu_i \\ & - \sum_k \mu_k \mathbb{1}_k^h \Lambda \frac{v_k}{v} + \sum_k \mu_k \mathbb{1}_{i\to k}^p \xi\Lambda \frac{v_k}{v} - \mu_i \xi\Lambda \sum_k \mathbb{1}_{i\to k}^p \frac{v_k}{v} \\ & + (1-\xi)(1-\theta)\frac{\Lambda}{v}\frac{1}{u+\xi(1-u)}\Bigg(\sum_k \mu_k \mathbb{1}_k^h v_k u \\ & + \xi \sum_k \mu_k v_k \left(\sum_j n_j \mathbb{1}_{j\to k}^p\right) - \xi \sum_k \mu_k n_k \left(\sum_j v_j \mathbb{1}_{k\to j}^p\right)\Bigg)\end{aligned}$$

where the first line corresponds to the instantaneous payoff of the marginal match, the second comes from the direct effect of n_i on the law of motion of $\dot{n}_i$, and the last two lines capture the indirect congestion effect through Λ. Rearranging yields (7).

For each number of operating firms N, I solve for $\{\mu_i\}$, $\{v_i\}$, and $\{n_i\}$ using an algorithm similar to Appendix D.3. I then compute the steady-state Hamiltonian:

$$\mathcal{H}^N = \left(\sum_i n_i z_i + \left(1 - \sum_i n_i\right) b - \sum_i c(v_i)\right) - c_f N$$

	(i)	(ii)	(iii)	(iv)	(v)	(vi)
HHI^{emp}	0.056	-0.266			0	
HHI^{wb}			0.060	-0.262		0
Prod.		2.02		2.03		
Fixed effects					Firm	Firm

Table 13: Wage regressions in the model with and without productivity control for different fixed effects

I repeat the exercise for each N, and choose the one maximizing $\mathcal{H}^N$.

D.5 Panel regressions through the lens of the model

I use the model to clarify how regression controls affect the estimated wage–concentration relationship. I vary the operating cost, simulate labor markets, and generate worker histories. Using the same definitions as in the data, I construct a synthetic panel and regress worker wages on market concentration, with and without controls for firm productivity or firm fixed effects. Table 13 reports the results.

Both concentration measures deliver similar patterns. Two results stand out. First, in this steady-state simulation, wages are time-invariant within firm, so firm fixed effects absorb essentially all wage variation; correspondingly, coefficients move toward zero when I include firm fixed effects, as in Table 11. Second, conditioning on firm productivity substantially attenuates the concentration coefficient, reflecting the distinction emphasized in the main text: productivity controls (or firm effects) strip out the output-sharing/selection component and move the regression closer to a "market-power" object, while specifications without such controls preserve the general-equilibrium selection channel. The fact that firm fixed effects soak up wage variation in this exercise does not mean concentration is irrelevant; rather, it affects wages through cross-firm differences that the firm effects intentionally

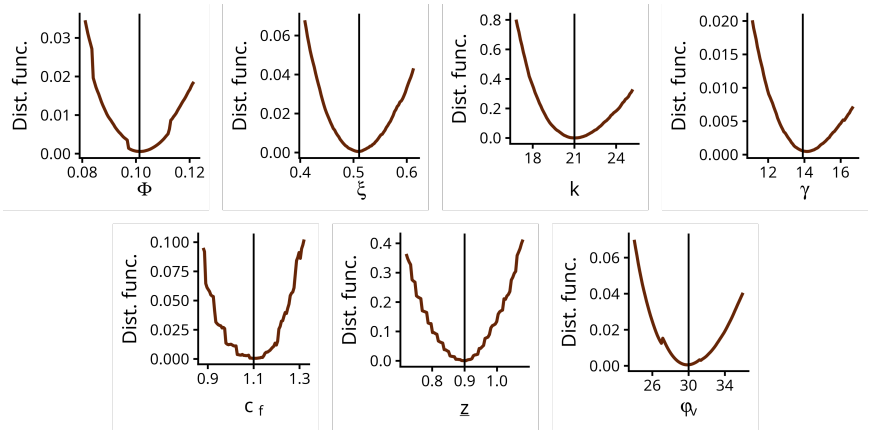


Figure 13: Sensitivity of the estimation

partial out.

E Robustness of the estimation

This appendix first looks at the sensitivity of the estimation, and then shows that the wage decomposition results are virtually unchanged compared to Section 4 when varying the matching elasticity θ, adjusting α to satisfy the Hosios condition in each case.

E.1 Sensitivity of the estimation

I study here the sensitivity of the estimation. I vary each parameter around its estimated value, holding the others fixed, and recompute the distance between data and model moments. Figure 13 plots this distance as a function of each parameter, with the vertical bar indicating the estimate. Each parameter is well estimated. I also study how each parameter affects its targeted moment. For each target I look at the distance between the data moment and its model counterpart as the associated parameter varies around its estimated value. The figure

is omitted for brevity, but parameters have a strong effect on their own target.

E.2 Robustness of the results to different values of θ and α

Tables 14 and 15 replicate the wage decomposition in Table 6 for alternative matching elasticities $\theta \in \{0.35, 0.65\}$, adjusting α accordingly to satisfy the Hosios condition, and re-estimating the other parameters. The decomposition shares and implied markdowns are very similar across specifications, although markdowns are higher when worker bargaining power is lower—as expected.

Operating cost c_f	0						
Percentile productivity	0%	10%	25%	50%	75%	90%	100%
Firm's productivity	0.81	0.82	0.83	0.86	0.92	0.99	1.32
Total wage	0.74	0.74	0.75	0.77	0.80	0.84	0.97
From output (%)	38.5	38.6	38.9	39.3	40.1	41.4	47.1
From penalty (%)	-8.3	-8.0	-7.5	-6.3	-4.4	-2.5	0.0
From outside option (%)	69.8	69.4	68.6	67.0	64.3	61.1	52.9
Markdown	0.09	0.09	0.10	0.11	0.13	0.15	0.26
Share workers (%)	0.1	0.1	0.2	0.3	0.7	1.7	4.4
HHI				0.019			

Operating cost c_f	1.68×10^{-3}						
Percentile productivity	0%	10%	25%	50%	75%	90%	100%
Firm's productivity	0.95	0.95	0.97	1.00	1.07	1.17	1.32
Total wage	0.78	0.78	0.79	0.81	0.85	0.90	0.96
From output (%)	42.7	42.8	42.9	43.3	44.1	45.7	48.1
From penalty (%)	-6.6	-6.2	-5.5	-4.2	-2.4	-0.9	0.0
From outside option (%)	63.8	63.4	62.7	60.9	58.4	55.3	51.9
Markdown	0.18	0.18	0.18	0.19	0.21	0.23	0.27
Share workers (%)	1.4	1.5	1.9	2.9	5.0	7.5	9.7
HHI				0.054			

Operating cost c_f	33.54×10^{-3}						
Percentile productivity	0%	10%	25%	50%	75%	90%	100%
Firm's productivity	1.18	1.18	1.18	1.22	1.31	1.32	1.32
Total wage	0.83	0.83	0.83	0.86	0.90	0.90	0.90
From output (%)	49.3	49.3	49.3	49.7	50.9	51.0	51.0
From penalty (%)	-2.1	-2.1	-2.1	-1.2	0.0	0.0	0.0
From outside option (%)	52.6	52.6	52.6	51.5	49.1	49.0	49.0
Markdown	0.29	0.29	0.29	0.30	0.31	0.31	0.31
Share workers (%)	26.8	26.8	26.8	32.6	40.8	40.8	40.8
HHI				0.349			

Table 14: Wage decomposition and markdowns by firm's productivity for $\theta = 0.35$ and $\alpha = 0.35$ (other parameters re-estimated)

Operating cost c_f	0						
Percentile productivity	0%	10%	25%	50%	75%	90%	100%
Firm's productivity	0.95	0.95	0.95	0.96	0.98	1.00	1.08
Total wage	0.87	0.88	0.89	0.90	0.93	0.95	1.02
From output (%)	70.5	70.2	69.9	69.3	68.7	68.4	69.4
From penalty (%)	-6.1	-5.7	-5.0	-3.8	-2.3	-1.1	0.0
From outside option (%)	35.7	35.5	35.1	34.5	33.6	32.7	30.6
Markdown	0.08	0.07	0.07	0.06	0.05	0.05	0.06
Share workers (%)	0.1	0.1	0.2	0.3	0.7	1.7	4.6
HHI				0.019			

Operating cost c_f	0.63×10^{-3}						
Percentile productivity	0%	10%	25%	50%	75%	90%	100%
Firm's productivity	0.99	0.99	0.99	1.00	1.02	1.05	1.08
Total wage	0.90	0.90	0.91	0.93	0.95	0.98	1.01
From output (%)	71.5	71.3	71.0	70.3	69.7	69.5	69.7
From penalty (%)	-5.5	-5.2	-4.5	-3.2	-1.7	-0.6	0.0
From outside option (%)	34.0	33.8	33.5	32.9	32.0	31.2	30.3
Markdown	0.09	0.09	0.08	0.08	0.07	0.07	0.07
Share workers (%)	0.9	1.0	1.3	2.3	4.9	9.3	14.3
HHI				0.071			

Operating cost c_f	12.60×10^{-3}						
Percentile productivity	0%	10%	25%	50%	75%	90%	100%
Firm's productivity	1.06	1.06	1.06	1.06	1.08	1.08	1.08
Total wage	0.95	0.95	0.95	0.96	0.99	0.99	0.99
From output (%)	72.4	72.4	72.4	72.1	71.0	71.0	71.0
From penalty (%)	-2.6	-2.6	-2.6	-2.2	0.0	0.0	0.0
From outside option (%)	30.4	30.4	30.4	30.0	29.0	29.0	29.0
Markdown	0.10	0.10	0.10	0.10	0.08	0.08	0.08
Share workers (%)	30.5	30.5	30.5	30.5	69.2	69.2	69.2
HHI				0.578			

Table 15: Wage decomposition and markdowns by firm's productivity for $\theta = 0.65$ and $\alpha = 0.65$ (other parameters re-estimated)

	Total output	Nbr. firms	HHI^{emp}	Unemp. rate	Total mass of vacancies	Share vacancies by best firm	J2J prob.
Decentralized	0.71	15	0.081	68.8%	0.0	12.6%	0.7%
No congestion internalization	+0.2%	15	0.087	68.6%	0.0	13.7%	0.7%
Planner	+0.6%	7	0.189	72.8%	0.0	25.5%	1.0%

Table 16: Decentralized labor markets and planner's solution for $\gamma = 0.5$

F Counterfactuals appendix

F.1 Inefficiencies when reducing γ

Table 16 repeats the decentralized-versus-planner comparison for $\gamma = 0.5$. As γ decreases, vacancy costs become less convex and the planner concentrates employment and vacancy creation more strongly among the most productive firms (in the limit $\gamma \to 0$, the efficient allocation would collapse to a single operating firm). At the same time, output gains remain modest because lowering γ scales up total vacancy-posting costs for a given level of vacancies. Finally, the quantitative role of firms internalizing their impact on aggregate vacancy congestion remains small even at low γ.

F.2 Second best policies at low and high concentrations

Table 17 repeats the second-best policy exercise from Section 5 in a low- and a high-concentration market. In both cases, an entry tax that deters low-productivity firms from operating delivers most of the efficiency gains by selecting the same set of operating firms as in the planner's allocation, raising output. Mean wages increase slightly, while mean markdowns also increase slightly as workers' outside option deteriorates. As the set

	Nbr. firms	HHI^{emp}	Unemp. rate	Total output	Total mass of vacancies	Mean markdown	Mean wage
Decentralized	154	0.018	9.2%	0.96	2.9	0.098	0.91
Entry tax	46	0.040	15.3%	+0.8%	1.0	0.120	0.92
Entry + average tax	46	0.040	15.8%	+0.9%	0.9	0.065	0.72
Planner	46	0.040	15.8%	+0.9%	0.9		

(a) Low concentration ($c_f = 0$)

	Nbr. firms	HHI^{emp}	Unemp. rate	Total output	Total mass of vacancies	Mean markdown	Mean wage
Decentralized	3	0.365	41.4%	0.85	0.1	0.179	0.93
Entry tax	2	0.528	46.5%	+0.1%	0.1	0.185	0.94
Entry + average tax	2	0.528	46.5%	+0.1%	0.1	0.184	0.93
Planner	2	0.527	46.5%	+0.1%	0.1		

(b) High concentration ($c_f = 22.00 \times 10^{-3}$)

Table 17: Decentralized labor markets, first-best, and second-best solutions at low and high concentration

of operating firms decreases (high-concentration markets), inefficiencies become smaller. Adding a production tax to target aggregate vacancy creation yields only small additional output gains and reduces mean wages mechanically by lowering the amount of output to share.