Monopsony Power and Firm Organization*

Álvaro Jáñez Lukas Delgado-Prieto

Universidad Carlos III de Madrid

July 10, 2023

Abstract

We develop a general equilibrium model where wage dispersion between managers and production workers stems from heterogeneous wage markdowns and firm organizational decisions. Quantifying the model to administrative data covering the universe of private firms in Portugal, we show that monopsony power reduces wage dispersion between occupations by 12 percent and generates welfare losses through distorting firms' managerial delegation decisions. The recent rise in the Portuguese minimum wage only mitigates the welfare loss from monopsony for managers. This policy increases wages for both occupations but only rises employment for managers, who benefit from the induced labor relocation of a portion of production workers toward large firms.

Keywords: Monopsony power, firm organization, welfare, minimum wages.

JEL: D21, J21, J31, J42, O40

---

*We are grateful to Felix Wellschmied and Jan Stuhler for their guidance and support, and to Antonia Díaz, Andrés Erosa, Luisa Fuster, Matthias Kredler, and Emircan Yurdagul for many insightful discussions. We are also thankful for comments from Manuel Amador, Timothy J. Kehoe, Illenin Kondo, Simon Mongey, and Juan Pablo Nicolini, as well as all the participants at the Ph.D. Workshop at Universidad Carlos III and the Trade Workshop at the University of Minnesota. We appreciate the invaluable effort made by João Pereira dos Santos and Marta Lopes to support our research project to access the data. Álvaro Jáñez gratefully acknowledges support from the Ministerio de Ciencia e Innovación through research grants *PRE2019-088620*, the financial support from the AEI (Spain) through research grant PID2019-110779GB-I00, and Universidad Carlos III through the scholarship for Master's study.

Email addresses: ajanez@eco.uc3m.es (Álvaro Jáñez), ludelgad@eco.uc3m.es (Lukas Delgado-Prieto).

1 Introduction

There is a close relationship between the number of competing firms in a local labor market and the internal organization of such firms. For instance, managerial job markets tend to have fewer participant firms relative to the markets of production workers, as only a fraction of firms delegate decision-making to managers. Thus, wage disparities between both worker types may also reflect differences in competition rather than only productivity. Simultaneously, the lack of labor market competition incentivizes low-productive firms to expand their organizational structure, leading to an inefficient allocation of managers and production workers across firms. This paper shows that understanding the interaction between firm organization and monopsony power helps us to explain wage dispersion across distinct types of workers, labor misallocation across firms, and the design of minimum wage policies.

To that end, we develop a general equilibrium oligopsony model where firms organize their production in layers and quantify it to Portuguese administrative data. Our main result shows that monopsony power reduces the dispersion in mean wages between occupations by 12 percent and leads to welfare losses for both occupations through the distortion of firms' managerial delegation decisions. We find that increasing the minimum wage fails to alleviate the welfare losses from monopsony power for production workers. On the contrary, the minimum wage reform benefits managers whose employment and wages rise because this policy induces the reallocation of a portion of production workers toward large firms.

To arrive at these conclusions, we first use data from an annual census of employees in Portugal. For each occupation-specific local labor market, we compute the Herfindahl-Hirschman Index (HHI) which represents a weighted average of firms' payroll shares. We document two stylized facts. First, the average manager works in a local labor market where the HHI is about 38 percent greater than in the market of the average production worker. This heterogeneity stems from the fact that the markets of production workers have about two times as many firms as the markets of managers and because a higher proportion of managers than production workers sort into markets with high levels of concentration. Second, highly concentrated labor markets have a higher share of firms that delegates the supervision of production workers to managers. In particular, an increase of 10 percentage points in the HHI is associated with an increase of 3 percentage points in the share of multi-layer firms. Overall, these facts point out an inherent relationship between the organizational decisions of firms and their degree of monopsony power.

To quantify the distribution of wage markdowns for the aforementioned occupations and perform counterfactual simulations, we develop a general equilibrium model that incorporates oligopsony competition, minimum wages, and firm organization. The economy features a representative household for each occupation and a continuum of locations, each with a finite number of firms. Households choose consumption and the measure of workers to supply to each firm for their respective occupation. They view firms and locations as imperfect substitutes, where the degree of substitutability is exogenously occupation-specific. Moreover, labor disutility costs are also exogenously heterogeneous across occupations. Firms exogenously differ in terms of productivity and the location they inhabit. The firm's decision consists of adopting either a single- or multi-layer organizational structure. Single-layer firms decide how many production workers to hire. Multi-layer firms include an additional managerial layer and choose how many managers and production workers to hire. In the model, managers enable firms to direct a larger workforce, yet it comes at higher overhead costs, and, thus, only the most productive firm adopt a multi-layer structure. Regarding monopsony power, imperfect firm substitutability, together with firms being non-atomistic within their location, imply that firms face upward-sloping labor supply curves and make employment choices under strategic competition, i.e., internalizing the employment policies of their competitors in the same location. When minimum wages do not constrain the employment and wage decisions of firms, this environment leads to equilibrium wage markdowns that depend on (i) the exogenous occupation-specific parameters determining how substitutable firms are across different (θ_o) and within the same (η_o) location from the household's perspective and (ii) the market payroll share of the firm, which endogenously depends on the organizational and employment choices of firms.

We estimate the model using the Simulated Method of Moments (SMM) to fit moments of firm organization, wages, and market concentration. We estimate the within-market elasticities from the correlation between employment and wages within each market. Moreover, we adopt an indirect inference approach to estimate the across-market elasticities from plausibly exogenous labor demand changes at the municipality level generated by national sector employment trends. We estimate that both elasticities are lower for managers, implying that it is more costly for managers to switch firms within the same location and across distinct locations. In addition to the lesser influence of minimum wages on managerial salaries and the fact that managers mostly work for large firms with high payroll shares, this result implies that firms generally exert wider wage markdowns on managers than on production workers' wages. On average, we estimate that firms markdown managerial wages by 27 percent and production worker wages by 11 percent.

To quantify the effect of monopsony power on aggregate outcomes, we compare the benchmark equilibrium with a counterfactual efficient economy where wages are equal to the marginal product of labor. First, we find that monopsony power reduces wage dispersion between managers and production workers, as managers earn higher average wages and bear wider markdowns. In particular, the mean managerial-to-production worker wage ratio would rise by 12 percent. Second, we quantify an output loss of 2.6 percent from monopsony power. Monopsony power leads to a reduction in aggregate labor supply because it lowers wages relative to the efficient economy. Furthermore, since the most productive firms restrict employment the most because they set wider wage markdowns, monopsony power also distorts the allocation of labor across firms, especially of managers. Third, we estimate a welfare gain of 5 percent for production workers and 24 percent for managers in the efficient economy, as both of them benefit from the efficiency gains and the redistribution of firm profits to labor income.

Lastly, we assess the effectiveness of recent Portuguese minimum wage reforms in improving the welfare of households by further limiting the monopsony power of firms. We simulate an increase of 10 percent in the real minimum wage, the same increase we observe in the data during the three years after our calibration period. The increase in the minimum wage has almost no effect on aggregate employment and efficiency. However, it has heterogeneous effects on employment and welfare across occupations. On the one hand, the reform decreases the employment level of production workers, as a big share of them work in low-productivity firms that pay wages close to the minimum wage and exert narrow markdowns. In addition, it increases their disutility of labor supply because the reforms induce a reallocation of labor across firms. Overall, these two channels crowd out the gains from higher wages and imply that the welfare of production workers decreases by 1.23 percent. On the other hand, the welfare of managers rises by 1.08 percent. The reason is that the reform indirectly increases both employment and wages for managers. Although the minimum wage is well-below most of their wage, managers benefit from the indirect effect of the reform through two channels. First, firms have more incentives to demand managers because their relative cost falls, as the reform especially increases the wage of production workers. Second, the reallocation of production workers toward high-productivity firms incentivizes these firms to hire new managers and enhance the productivity of the incumbent managers.

Literature. This paper relates to the empirical literature on monopsony and labor market concentration (Martins, 2018; Azar et al., 2020; Benmelech et al., 2020; Rinz, 2022; Azar et al., 2022; Bassanini et al., 2023; Dodini et al., 2023). We bring two new stylized facts to this literature. First, we document that managerial markets display higher levels of market payroll concentration relative to the markets of production workers. Second, we highlight that this heterogeneity stems from the fact that managerial markets have fewer firms and because managers sort into more concentrated markets.

We also contribute to the literature on oligopsony (Bhaskar et al., 2002; MacKenzie, 2021; Berger et al., 2022, 2023a; Jarosch et al., 2023). Using models where labor market power arises from firm granularity and imperfect firm substitutability, they study the effect of labor market power on wages, efficiency, and welfare. Our main contribution to this literature is to study the effect of labor market power on these outcomes through the organization of work within firms. In our model, firms engage in organizational decisions that give rise to wage dispersion and markdown heterogeneity for production workers and managers.

We connect to the literature that studies the effect of minimum wage policies in models with imperfect labor market competition (Bamford, 2021; Ahlfeldt et al., 2022; Hurst et al., 2022; Karabarbounis et al., 2022; Drechsel-Grau, 2023). Our framework is particularly close to Berger et al. (2023b), who studies the effect of minimum wages on efficiency and welfare in an oligopsonistic environment with firm and worker heterogeneity. By considering firm organizational decisions, we add two contributions to this literature. First, we show that rising the minimum wage increases the share of firms that delegate on managers and reduces the managerial span of control. Second, we show that rising minimum wage policies differently affect managers and production workers.

This paper contributes to the literature on firm production organization, which rationalizes organizations as an efficient device to use and communicate the knowledge required for production (Garicano, 2000; Garicano and Rossi-Hansberg, 2006; Caliendo and Rossi-Hansberg, 2012). Several studies build on this model to analyze firm-size distortions (Garicano et al., 2016; Tamkoç, 2022), the adoption of information and technological capital Mariscal (2020), the misallocation of labor in developing countries (Grobovsek, 2020), and technological adoptions across urban areas (Santamaria, 2023). To our knowledge, we are the first to combine imperfect labor market competition and firm organization in a general equilibrium model to rationalize differences in monopsony power.

Lastly, we relate to the literature that studies the misallocation of labor across firms (Hsieh and Klenow, 2009; Bartelsman et al., 2013; Davis et al., 2014; Garcia-Santana and Pijoan-Mas, 2014; Heise and Porzio, 2023). We show that large firms especially restrict managerial employment relative to the efficient level in Portugal, which leads to an inefficiently low share of multi-layer firms and inefficiently high managerial span of control.

2 Stylized Facts

This section documents two stylized facts that link the organizational decisions of firms with market payroll concentration. First, the average manager works in a local labor market with a Herfindahl-Hirschman Index, a standard market concentration measure, 38 percent greater than the average production worker. Second, the share of firms that delegate the supervision of production workers to managers increases with market concentration.

2.1 Data

Our primary data source is *Quadros de Pessoal* (QP), an annual census of private sector employees conducted by the Portuguese Ministry of Employment. This census provides matched employer-employee data with information on employment, monthly wages, occupation, industry, and municipality for all firms based in Portugal with at least one worker. Our sample period covers from 2002 to 2016.

We classify labor markets based on three observable characteristics of the job: geography, industry, and occupation. This classification stems from the fact that workers are more attached to their current labor market because of the imperfect substitutability of skills across jobs and sectors, as well as imperfect geographical mobility (Neal, 1995; Kambourov and Manovskii, 2009; Sullivan, 2010; Kennan and Walker, 2011; Monte et al., 2018). In particular, we define two broad occupations, i.e., managers and production workers, and define a local labor market for each occupation as the intersection of the geography (municipality) and industry (2-digit NACE).

We assign workers to each occupation following a hierarchical classification similar to Caliendo et al. (2020). Particularly, we partition professional categories into two layers. We consider “top executives", “intermediary executives", and “supervisors and team leaders" as managers, while we group the remaining categories as production workers. In the data, the critical difference between both occupations is that managers are responsible for the organizational policies of the firm and their adaptation, which require a high degree of qualification in

Table 1: Market Concentration by Occupation

	Managers	**Production Workers**
	Mean	Mean
Max s_{ij}	0.38	0.30
HHI_j	0.27	0.20

Source: elaboration based on Quadros de Pessoal.

Note: The Table reports the employment-weighted mean and median of the HHI and maximum payroll share across local labor markets.

terms of direction, guidance, and coordination of the firm fundamental activities.[1]

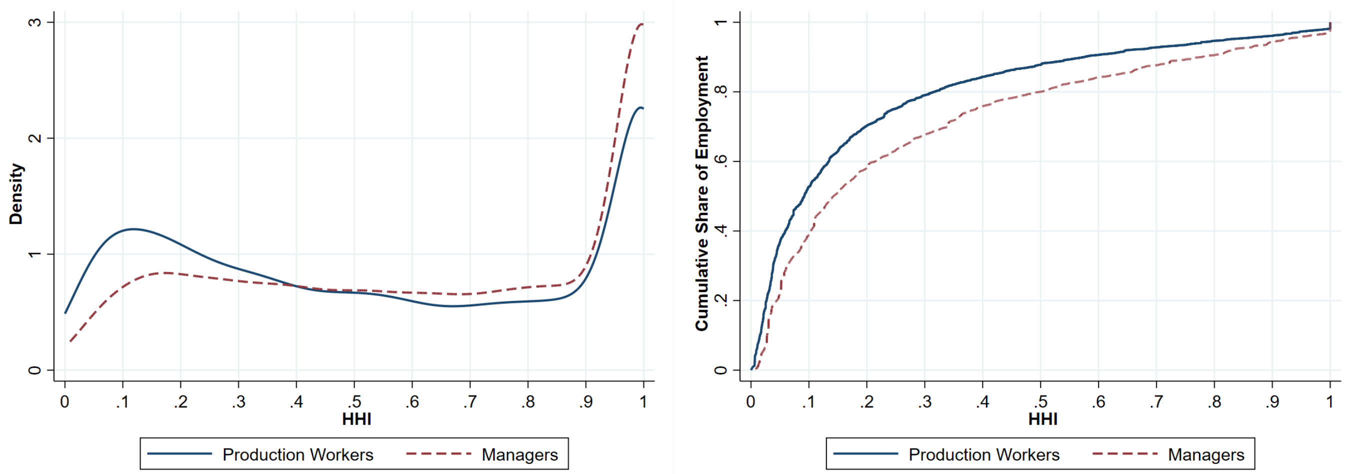
2.2 Payroll Concentration and Firm Organization

The level of market payroll concentration in a local labor market, which represents how much of the total market payroll is in the hands of a few firms, is a common proxy for the degree of monopsony power that firms hold in such a market. We measure market payroll concentration in each local labor market using the Herfindahl-Hirschman Index (HHI). This index is equal to a weighted average payroll share of firms within the market. Thus, an increase in the HHI reflects higher market concentration because fewer firms accumulate a greater share of the market payroll.

Fact #1: Managers work in more concentrated markets. In Table 1, we document that the average manager works in a market whose HHI is 38 percent higher than the average production worker. Specifically, the employment-weighted average HHI is 0.27 for managers and 0.20 for production workers. To provide context for these numbers, one would observe the same level of concentration with three and five equally-sized establishments, respectively. Table 1 also reports the employment-weighted average of the maximum payroll share across local labor markets. The average largest establishment in a local labor market accumulates

[1]See Table A.1 for further information about the categories of the occupational classification, which is based on *Decreto-Lei nº. 121/78 de 2 de Junho, Ministério do Trabalho*. Figure A.2 shows that most workers remain within the same category after changing to other firms.

Figure 1: Distribution of HHI by Occupation



(a) Market Concentration

(b) Cumulative Employment

Source: Elaboration based on QP.
Note: The Graph plots the kernel density of the payroll HHI across local labor markets for managers (red) and production workers (blue). In addition, the graph plots the cumulative share of employment with respect to the payroll HHI.

about eight percentage points more market payroll in the market of managers than in the market of production workers.

Two mechanisms explain why managers tend to work in more concentrated markets than production workers. First, heterogeneity in the distribution of payroll concentration across markets. Second, heterogeneity in employment sorting across these market types. We display both channels in Figure 1, which plots the estimated kernel density of the market level HHI for each occupation and their respective cumulative share of employment. We find that managers work in more concentrated markets because (i) their markets are more likely to present higher payroll concentration (see Figure 1a) and because (ii) there is a higher share of managers working in such markets (see Figure 1b).

Regarding the first observation, we next use a standard decomposition that breaks the HHI into two elements to understand what makes managerial markets more likely to be concentrated (see Appendix B.4). The first element involves the *number of establishments* in each

Table 2: Decomposition of the Average HHI across Occupations

	Component: Nº establishments	Component: Dispersion in shares	Mean HHI
(1) Managers	0.57	0.08	0.65
(2) Production Workers	0.47	0.09	0.56
(1)-(2) Gap	0.10	-0.01	0.09

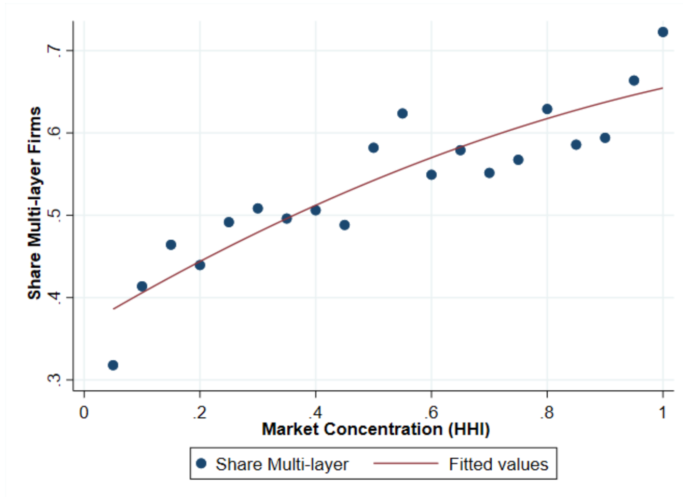
Note: The table reports the contribution of each channel to the unweighted average HHI level across occupations. The second column reports the contribution of the number of establishments ($1/N$), while the third column shows the contribution of the dispersion in payroll shares relative to the symmetric case $\left(\sum_{i=1}^{N}\left(s_i - \frac{1}{N}\right)^2\right)$.

market. All else being constant, increasing the number of firms lowers the size of firms in the market. The second element entails the *dispersion level of payroll shares* across establishments relative to the case in which they hold identical shares. All else being constant, increasing the dispersion in payroll shares leads to a greater share of market payroll in fewer firms. Table 2 reports the results from decomposing the unweighted average of the HHI for both occupations. We do not weigh each market by employment size to isolate the decomposition from employment sorting. We find that the entire gap of nine percentage points stems from the fact that managerial markets have fewer firms. In particular, the markets of production workers tend to have almost two times as many establishments as the markets of managers (see Table A.6). This fact emphasizes the importance of modeling the decision-making of firms regarding managerial delegation, as only a proportion of establishments find it necessary to hire managers to organize their production.

The second observation emphasizes the need to model the allocation of labor across markets. Overall, most employees in both occupations work in a handful of markets with relatively low concentration levels. However, the proportion of managers in those markets is comparatively smaller. For instance, nearly 70 percent of production workers are in markets with a HHI below 0.20, whereas only 60 percent of managers work in such markets.

Fact 2#: Higher share of multi-layer firms in more concentrated markets. Figure 2 displays the average share of firms that hire both managers and production workers across

Figure 2: Market Concentration and Multi-layer Firms



Source: Elaboration based on QP.
Note: The Graph plots the average share of multi-layer firms across local labor markets with different levels of concentration. For each local labor market of production workers, we compute the share of multi layer firms and the HHI. We split the distribution of the HHI into 20 cells of length 0.05. In each cell, we take the employment weighted mean of the share of multi-layer firms across markets.

markets with different levels of the HHI. When labor market concentration is higher, there is a higher share of firms that delegates supervision of production workers to managers. Rising the HHI by 10 percentage points is associated with an increase of about 1 percentage points in the share of multi-layer firms. This suggests a tight relationship between the internal organization of firms and the level of payroll concentration in the labor markets where they operate. Two complementary stories that have different economic implications may explain this result. On the one hand, this relationship may reflect the tendency of high-productivity firms to adopt a multi-layer structure as they expand, consequently contributing to increased market concentration. On the other hand, it may reflect that small markets with low competition allow, even low-productivity firms, to increase their size and hire managers because mobility and search frictions prevent production workers from relocating to other firms.

Firm organization and monopsony power. To sum up, the internal organization of firms is closely related to market concentration and, thus, potentially also to monopsony power. We stand out two reasons to analyze this relationship through the lens of a general equilibrium model. First, the fact that managerial markets are more concentrated suggests that firms have greater monopsony power over managerial wages. Yet, this market categorizations do not capture the entire set of potential employers, as workers flow across industries and regions. Thus, to quantify monopsony power, it is necessary to measure the extent to which workers find it costly to switch firms both within the same and across different markets. Second, the fact that more concentrated markets have a higher share of multi-layer firms suggests a relationship between monopsony power and the organization of work within firms. Thus, monopsony power may affect aggregate output through the organizational choices of firms, and policies aimed at curbing monopsony power may also alter the structure of firms.

3 Model

This section presents a general equilibrium model that incorporates managerial delegation, oligopsonistic labor markets, and minimum wages. The model considers two occupations, managers and production workers, each with exogenously heterogeneous labor disutility costs and firm substitutability parameters. For each occupation, there is a household that makes consumption choices and decides the labor supply to each firm. Firms are exogenously heterogeneous in productivity and the location they inhabit. Regarding organization, firms choose whether to delegate production workers to managers. Conditional on the organization structure, they choose the number of workers in each occupation. When making employment choices, firms have monopsony power and face a minimum wage. The oligopsony structure implies that unconstrained firms exert occupation-specific wage markdowns, which depend on their market payroll share and parameters determining firm substitutability from the worker perspective. Since the distribution of payroll shares arises from the organizational choices of firms, the model endogenously captures differences in wage markdowns across both occupations.

3.1 Environment

Agents. The economy is populated by two households, indexed by their occupation type $o \in \{w, m\}$, and a continuum of firms. Households are exogenously heterogeneous in terms of disutility costs for each unit of work and how substitutable they view different firms. Firms are exogenously heterogeneous in two dimensions. First, they belong to a continuum of locations $j \in [0, 1]$, where each location contains a finite number of firms indexed by $i \in \{1, \dots, M_j\}$. Note that each location contains a local labor market for each occupation, i.e., the data counterpart of the location is the intersection between region and industry. Second, firms differ in productivity z_{ij}, which is drawn from a standard log-normal distribution with standard deviation σ_z.

Goods and Technology. Firms use labor to produce a tradable good in a perfectly competitive national market whose price we normalize to one. We assume that there are two types of labor: production workers and managers. While production workers are essential for production, managers are optional. We assume each firm chooses between two types of organizations which vary in the number of layers, $\ell \in \{1, 2\}$.

The first type of firm organization is the single-layer firm ($\ell = 1$). Single-layer firms use n_w units of labor of production workers to produce according to the following technology:

$$y(z, 1) = z \, \varphi_w \, n_w^{1-\alpha}. \tag{1}$$

The parameter φ_w stands for the efficiency of labor of production workers. This technological specification includes the essential features of the production technology in the standard model of Lucas (1978). Namely, firms are heterogeneous in productivity and face diminishing returns to scale. The parameter $1-\alpha \in (0, 1)$ determines the strength of diminishing returns and, thus, limits the span of control of single-layer firms.

Alternatively, firms may choose to be multi-layer organizations ($\ell = 2$) in order to manage a larger workforce. In this case, firms additionally include a managerial layer and use the labor of both managers n_m and production workers n_w to produce output. The technology of multi-layer firms is given by:

$$y(z,2) = z\,\varphi_m \varphi_w\, n_m^{(1-\alpha)\,\alpha} n_w^{1-\alpha}. \tag{2}$$

The parameter φ_m reflects how much managers enhance the productivity of production workers. Under this technological specification, managerial delegation allows firms to increase their workforce by dampening the diminishing returns to scale, as managers increase the marginal productivity of production workers. In addition, the parameter α embeds in a simple manner all the technical reasons that influence the span of control of managers, as it is the only parameter that is informative of the ratio of the marginal labor productivity between the two occupations.

Overall, this technological specification abstracts from the micro-foundations in the theory of firm organization, such as those of the knowledge-based hierarchy literature (Garicano, 2000; Garicano and Rossi-Hansberg, 2006). However, it allows for the quantification of the model while having the main organizational trade-off: adding a managerial layer permits the firm to increase its size, but it comes at the expense of higher costs.[2]

Households. Each household type $o \in \{w, m\}$ chooses the measure of workers to supply to each firm n_{ijo} and consumption of each good c_{ij} to maximize their utility:

$$\mathcal{U}_o = \max_{n_{ijo}, c_{ijo}} \quad \mathbf{C}_o \; - \; \phi_o \, \frac{\mathbf{N}_o^{\,1+\frac{1}{\gamma}}}{1+\frac{1}{\gamma}}, \tag{3}$$

subject to the household's budget constraint:

$$\mathbf{C}_o = \int_0^1 \sum_{i=1}^{M_j} w_{ijo} n_{ijo} \, dj + \kappa_o \Pi, \tag{4}$$

[2] In an environment with heterogeneous firms endowed with monopsony power, adding a knowledge choice considerably constrains the parameter space for which an equilibrium exists. The intuition is that the knowledge choice of firms is inherently linked to their wage policy, making one of the two redundant. Then, the vector of wages that clears the market likely makes a significant proportion of firms choose zero knowledge, i.e., they optimally choose self-employment, making firm organization unimportant in the economy.

where we define the aggregate consumption and labor supply indexes as

$$\mathbf{C}_o := \int_0^1 \sum_{i=1}^{M_j} c_{ijo}\, dj$$

$$\mathbf{N_o} := \left[\int_0^1 \left(\frac{\mathbf{n_{jo}}}{B_{jo}} \right)^{\frac{\theta_o+1}{\theta_o}} dj \right]^{\frac{\theta_o}{\theta_o+1}} \quad \mathbf{n_{jo}} := \left[\sum_{i=1}^{M_j} n_{ijo}^{\frac{\eta_o+1}{\eta_o}} \right]^{\frac{\eta_o}{\eta_o+1}}, \quad \eta_o > \theta_o > 0.$$

Note that each household type gets a share κ_o of total firm profits. The parameter γ stands for the aggregate Frisch elasticity of households, ϕ_o is an occupational-specific labor disutility shifter, and B_{jo} is a market amenity shifter. We follow closely the consumption and labor supply structure of Berger et al. (2022). That is, we assume that consumption goods are perfectly substitutable, but households view firms as imperfect substitutes in terms of non-wage characteristics. In particular, each firm faces an occupation-specific upward-sloping labor supply curve with two elasticities of substitution $\theta_o > 0$ and $\eta_o > 0$. The parameter θ_o regulates the degree of substitutability of firms in distinct markets and, thus, captures the costs of moving across locations or idiosyncratic tastes for the location. When these costs decrease ($\theta_o \uparrow$), workers find it easier to substitute firms across locations and become more responsive to market wage differentials. The parameter η_o regulates the degree of substitutability of firms within the same market and, thus, captures features such as commuting costs, search costs, or idiosyncratic tastes for the firm. When these costs decrease ($\eta_o \uparrow$), workers find within-market, across-firm substitutability easier and become more responsive to wage differentials across firms in the same market. We refer to η_o and θ_o as the within- and across-market firm substitutability parameters. Critical to the main conclusions of the theory, we impose the following assumption:

Assumption 1. Across-market mobility costs are higher than within-market mobility costs:

$$\eta_0 \geq \theta_0, \ \forall o \in \{0, 1\} \tag{5}$$

In other words, we assume that both household types find firms within the same market as closer substitutes than firms in different markets. Under this assumption, larger firms hinder the relocation of their workers to other firms because their workers have fewer alternatives within the same market and, thus, need to move to other markets to find more potential employers. As a result, they become less responsive to the wage policy of the firm.

Linear utility in consumption implies that the aggregate disutility of labor supply in occupation o is:

$$\mathbf{N}_o = \left(\frac{\mathbf{W}_o}{\phi_o}\right)^{\gamma}, \tag{6}$$

and the labor supply curve of occupation o to firm i in market j is:

$$n_{ijo} = B_{jo}^{1+\theta_o}\left(\frac{w_{ijo}}{\mathbf{w}_{jo}}\right)^{\eta_o}\left(\frac{\mathbf{w}_{jo}}{\mathbf{W}_o}\right)^{\theta_o}\mathbf{N}_o \quad \longleftrightarrow \quad \underbrace{w_{ijo} = \left(\frac{1}{B_{jo}}\right)^{\frac{1+\theta_o}{\theta_o}}\left(\frac{n_{ijo}}{\mathbf{n}_{jo}}\right)^{\frac{1}{\eta_o}}\left(\frac{\mathbf{n}_{jo}}{\mathbf{N}_o}\right)^{\frac{1}{\theta_o}}\mathbf{W_o},}_{\text{Inverse labor supply curve } \forall ijo} \tag{7}$$

where we define the market wage index $\mathbf{w}_{jo}$ and the aggregate wage index $\mathbf{W}_o$ as

$$\mathbf{w}_{jo} := \left[\sum_{i\in j} w_{ijo}^{1+\eta_o}\right]^{\frac{1}{1+\eta_o}} \qquad \mathbf{W_o} := \left[\int_0^1 \left(B_{jo}\mathbf{w_{jo}}\right) z^{1+\theta_o} dj\right]^{\frac{1}{1+\theta_o}}. \tag{8}$$

We derive the labor supply system of equations given by (6) and (7) in Appendix C.1.

Firms. Firms choose the organizational structure to maximize profits, which consists of choosing whether to add a managerial layer:

$$\pi(z) = \max_{\ell}\ \{\pi(z,\ell)\}_{\ell=1}^{2}, \tag{9}$$

Conditional on the organizational structure, firms are constrained by the minimum wage, and we assume an oligopsonistic framework where they face upward-sloping labor supply curves and internalize the effect of their decisions on local market outcomes. Infinitesimal with respect to the economy, firms take the aggregate disutility of labor supply $\mathbf{N}_o$ and aggregate wages $\mathbf{W}_o$ as given. However, non-atomistic with respect to the local market, each firm internalizes the impact of the employment decisions of all firms in the market, including itself, on the labor supply curve it faces.

Single-layer firms choose the measure of production workers n_{ijw} to maximize profits, given the employment policies of their local competitors, n^*_{-ijw}. In particular, they solve:

$$\pi(z,1) \;=\; \max_{n_{ijw}} \quad y(z,1) \;-\; w_{ijw}\Big(n_{ijw}, n^*_{-ijw}, \mathbf{N}_w, \mathbf{W}_w\Big)n_{ijw}, \tag{10}$$

subject to the inverse labor supply curve of production workers and minimum wages:

$$w_{ijw}\Big(n_{ijw}, n^*_{-ijw}, \mathbf{N}_w, \mathbf{W}_w\Big) = \left(\frac{1}{B_{jo}}\right)^{\frac{1+\theta_o}{\theta_o}} \left(\frac{n_{ijw}}{\mathbf{n}_{jw}(n_{ijw}, n^*_{-ijw})}\right)^{\frac{1}{\eta_w}} \left(\frac{\mathbf{n}_{jw}(n_{ijw}, n^*_{-ijw})}{\mathbf{N}_w}\right)^{\frac{1}{\theta_w}} \mathbf{W}_w,$$

$$\mathbf{n}_{jw}(n_{ijw}, n^*_{-ijw}) = \left[n_{ijw}^{\frac{1+\eta_w}{\eta_w}} + \sum_{k\neq i} n^*_{kjw}{}^{\frac{1+\eta_w}{\eta_w}}\right]^{\frac{\eta_w}{1+\eta_w}},,$$

$$w_{ijw} \geq \underline{w}$$

Multi-layer firms choose the measure of production workers n_{ijw} and managers n_{ijm} to maximize profits, given the employment policies of their local competitors, (n^*_{-ijw}, n^*_{-ijm}). Their maximization problem is:

$$\pi(z,2) \;=\; \max_{n_{ijw}, n_{ijm}} \quad y(z,2) \;-\; \sum_{o\in\{w,m\}} w_{ijo}\Big(n_{ijo}, n^*_{-ijo}, \mathbf{N}_o, \mathbf{W}_o\Big)n_{ijo}, \tag{11}$$

subject to the inverse labor supply curve of both occupations:

$$w_{ijo}\Big(n_{ijo}, n^*_{-ijo}, \mathbf{N}_o, \mathbf{W}_o\Big) = \left(\frac{1}{B_{jo}}\right)^{\frac{1+\theta_o}{\theta_o}} \left(\frac{n_{ijo}}{\mathbf{n}_{jo}(n_{ijo}, n^*_{-ijo})}\right)^{\frac{1}{\eta_o}} \left(\frac{\mathbf{n}_{jo}(n_{ijo}, n^*_{-ijo})}{\mathbf{N}_o}\right)^{\frac{1}{\theta_o}} \mathbf{W}_o,$$

$$\mathbf{n}_{jo}(n_{ijo}, n^*_{-ijo}) = \left[n_{ijo}^{\frac{1+\eta_o}{\eta_o}} + \sum_{k\neq i} n^*_{kjo}{}^{\frac{1+\eta_o}{\eta_o}}\right]^{\frac{\eta_o}{1+\eta_o}},$$

$$w_{ijw} \geq \underline{w}, \quad \forall o \in \{w,m\}.$$

Given the number of the layers $\ell \in \{1, 2\}$ and occupation $o \in \{w, m\}$, the solution to the problem of the firm has three cases. First, firms are unconstrained by the minimum wage. Second, firms are constrained by the minimum wage, and labor demand is on the labor supply curve. In the aforementioned two cases, labor demand equals labor supply. Third, firms are constrained by the minimum wage, and labor demand is off the labor supply curve. In this case, there is an excess of labor supply at the minimum wage because the optimal wage is above the efficient level. We summarize the system of first-order conditions for each case as follows.

Case I: Minimum wage is not binding. Firms choose the level of employment in occupation $o \in \{w, m\}$ for which the marginal cost of labor equals its marginal product. Since firms are unconstrained, the optimal wage is above the minimum wage.

$$w^*_{ijo} = \mu_{ijo} \left.\frac{\partial y(z,\ell)}{\partial n_{ijo}}\right|_{n^*_{ijo}}, \qquad \mu_{ijo} = \frac{\varepsilon_{ijo}}{\varepsilon_{ijo}+1} \in (0,1), \qquad \varepsilon_{ijo} = \left[\frac{\partial \log w_{ijo}}{\partial \log n_{ijo}}\right]^{-1}. \tag{12}$$

Equation (12) determines that the marginal productivity of labor is higher than the wage payment for each occupation (see Appendix C.2 for complete derivations). As in the classical monopsony environment (Manning, 2013), the marginal cost of labor is equal to both the wage and the additional cost of increasing wages because firms internalize upward-sloping labor supply curves. This results in a wedge between wages and the marginal product of labor $\mu_{ijo} < 1$. In addition, firm granularity within each local labor market implies that wage markdowns and the labor supply elasticity are heterogeneous across firms because these internalize the impact of their relative market size on the labor supply curve they face. In particular, the structural labor supply elasticity has a closed-form solution and is given by:

$$\varepsilon_{ijo}(s_{ijo}) = \left[\frac{1}{\eta_o} + \left(\frac{1}{\theta_o} - \frac{1}{\eta_o}\right)\frac{\partial \log \mathbf{n}_{jo}}{\partial \log n_{ijo}}\right]^{-1} = \left[\frac{1}{\eta_o} + \left(\frac{1}{\theta_o} - \frac{1}{\eta_o}\right)s_{ijo}\right]^{-1}, \tag{13}$$

where s_{ijo} stands for the payroll share of firm i in market j:

$$s_{ijo} := \frac{w_{ijo} n_{ijo}}{\sum_{i \in j} w_{ijo} n_{ijo}}. \tag{14}$$

Note that firms face an elasticity that is a function of their payroll share for each occupation. Since we assume $\eta_0 \geq \theta_0$, larger firms face lower labor supply elasticities and exert wider markdowns. On the one hand, consider the monopsony case with only one firm in the market, i.e., $s_{ijo} = 1$. When a monopsonist makes an additional hire, it understands that it has to attract workers from other markets. Hence, it faces the across-market labor supply elasticity θ_{ijo}. On the other hand, consider an atomistic firm, i.e., $s_{ijo} \simeq 0$. This firm internalizes that a marginal hire attracts workers from the same market, as its decision has a negligible effect on market employment. Therefore, it faces the within-market labor supply elasticity η_{ijo}.

It is worth stressing that the model explicitly tells apart the forces shaping wage dispersion across occupations when firms are unconstrained. The first source of dispersion comes from differences in marginal productivity. The second source of dispersion comes from differences in markdowns. Firms have occupational-specific monopsony power because they may have different sizes in each local labor market (s_{ijo}) and because firm substitutability may differ across occupations (η_o, θ_o).

Case II: Minimum wage is binding and the firm is on the labor supply curve. Firms are constrained by the minimum wage, which is below the efficient wage level where the labor supply curve equals the marginal product of labor. In this case, the optimal wage is the minimum wage, the markdown is the ratio between the minimum wage and the marginal product, and the level of employment is given by the labor supply curve evaluated at the minimum wage.

$$w^*_{ijo} \;=\; \underline{w}, \qquad \mu_{ijo} \;=\; \frac{\underline{w}}{\frac{\partial y(z,\ell)}{\partial n_{ijo}}\Big|_{n^*_{ijo}}}, \qquad n^*_{ijo} = \left(\frac{\underline{w}}{\mathbf{w}_{jo}}\right)^{\eta_o}\left(\frac{\mathbf{w}_{jo}}{\mathbf{W}_o}\right)^{\theta_o}\mathbf{N}_o. \tag{15}$$

In this region, firms pay higher wages and hire more workers than they had done without the minimum wage.

Case III: Minimum wage is binding and the firm is off the labor supply curve. Firms are constrained by the minimum wage, which is above the efficient wage level where the labor supply curve equals the marginal product of labor. In this case, the optimal wage is equal to both the minimum wage and marginal product, the employment level is given by the marginal product, and the firm faces an excess of labor supply.

$$w^*_{ijo} = \underline{w} = \left.\frac{\partial y(z,\ell)}{\partial n_{ijo}}\right|_{n^*_{ijo}}, \qquad \mu_{ijo} = 1, \qquad n^*_{ijo} < \left(\frac{\underline{w}}{\mathbf{w}_{jo}}\right)^{\eta_o}\left(\frac{\mathbf{w}_{jo}}{\mathbf{W}_o}\right)^{\theta_o}\mathbf{N}_o. \tag{16}$$

In this region, firms pay higher wages and hire fewer workers than they had done without the minimum wage.

Equilibrium. Given a minimum wage $\underline{w}$, the general equilibrium of this economy is a set of organizational structures $\{\ell^*_{ij}\}$, aggregate disutilities of labor supply (N^*_w, N^*_m), and employment levels $\{n^*_{ijw}, n^*_{ijm}\}$ such that:

1. *Labor supply*: Households choose aggregate disutility N^*_o and labor supply to each firm $\{n^*_{ijo}\}$ to maximize utility. That is, Equation (6) and Equation (7) hold $\forall o \in \{w, m\}$.

2. *Firm organization*: Firms optimally choose the organizational structure: ℓ^*_{ij}. That is, Equation (9) holds $\forall j \in [0,1]$, $\forall i = \{1,\ldots,M_j\}$.

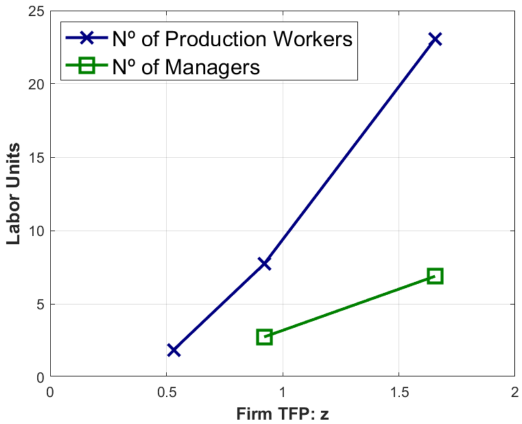
3. *Labor Demand*: Firms optimally choose employment (n^*_{ij0}, n^*_{ij1}). That is, Equations (12)-(16) hold $\forall j \in [0,1]$, $\forall i = \{1,\ldots,M_j\}$.

4. *Labor Market*: Labor supply and demand are equal for firms in Cases I and II. For firms in Case III, excess of labor supply is given by Equation (16).

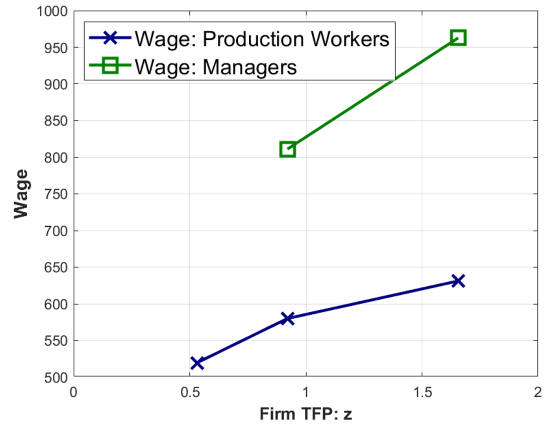
3.2 Market Characterization

Before presenting the quantification strategy of model parameters, we show the interaction between the firm organization, market concentration, and wage markdowns in equilibrium. Figure 3 plots the equilibrium level of employment, wages, and wage markdowns in a market with three firms. The numerical example displays the equilibrium in the model under the benchmark quantification.

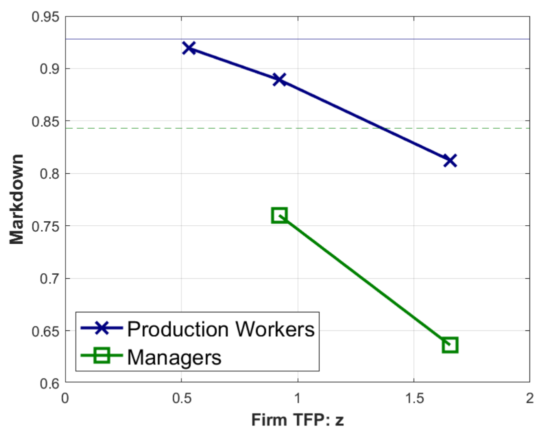
Figure 3: Oligopsonistic Market with Firm Organization



(a) Employment



(b) Wages



(c) Markdown

Note: Figure constructed from the model under the estimated parameters in Table 3.

Consider first the top left panel, which displays the employment level of production workers (blue) and managers (green) across firms. All firms hire production workers, as they are essential for production. However, only relatively high-productivity firms find it optimal to hire managers. These firms have more incentives to produce at large scales because of the technological complementarities between a firm's TFP and managerial labor. As a result, they use managerial delegation to dampen the diminishing marginal returns to the labor of production workers. The top right panel shows the wage level across firms for each occupation. Managers earn substantially higher wages than production workers, even conditional on the firm type. This occurs because the ratio of production workers to managers is sufficiently high, which drives up the marginal productivity of the latter with respect to that of the former. Moreover, the model captures the fact that wage dispersion between both occupations increases with firm size. This result stems from the fact that the ratio of managerial to production worker employment increases with firm size and therefore also the ratio of their marginal productivity.

Lastly, the bottom panel of Figure 3 plots wage markdowns across firms as the ratio between wages and the marginal product of labor for each occupation. Overall, wage markdowns are wider for managers than for production workers for two reasons. First, managers have a lower degree of firm-substitutability. If firms had the same payroll share in both markets, they would exert a wider markdown over managerial wages because they would internalize that managers bear higher across-firm mobility costs, which makes them less responsive to labor demand changes. Second, market payroll concentration is also higher for managers. If both occupations had the same degree of firm-substitutability, firms would exert a wider markdown over managerial wages because the labor supply elasticity of managers that firms face would be closer to the across-market elasticity than that of production workers.

4 Quantification of the Model

The quantification of the model parameters proceeds in two steps. Firstly, we exogenously calibrate the aggregate Frisch elasticity of labor supply and the profit share of each household type. Moreover, we endogenously calibrate the firm distribution and the market amenity shifter. Secondly, we estimate the remaining model parameters using the Simulated Method of Moments (SMM) approach. The quantification of the parameters targets moments of the firm organization, such as the average firm size or span of control. Together these moments determine the characteristics of the average firm across markets. Then, we target observed market differences in wages, employment, and payroll concentration. Moreover, we use an indirect inference approach to estimate the firm substitutability parameters from plausible exogenous changes in firms' labor demand.

4.1 Estimation

Table 3 summarizes the quantification of model parameters. We calibrate outside the model the aggregate Frisch elasticity and the share of profits related to each household type. To calibrate the aggregate Frisch elasticity, we follow Berger et al. (2022) by setting $\gamma = 0.5$, which is within the range that the Congressional Budget Office considers for policy evaluation. Lacking information about dividend payments across occupations, we set the share of profits of each household to their relative size in the economy. This implies that $\kappa_w = 0.81$ and $\kappa_m = 0.19$.

We calibrate inside the model the distribution of the number of firms across locations and market amenities. Regarding the firm distribution, $M_j \sim G(\cdot)$, we combine a discrete mass at $m_j = 1$ with a Pareto distribution to target the mass of markets with a single firm, the mean, and the standard deviation of the distribution of firms. Regarding market amenities, we note that only 12% of production workers belong to markets with less than ten firms, despite these markets representing nearly 65% of the total. We set a common market amenity for production workers in markets with less than ten firms, B_{ijw} such that $M_j \leq 10$, to target the share of production workers in those markets. The rationale for using amenities rather than productivity differences is that we would require too low productivity in small locations, which would highly overestimate the share of minimum wage earners in such places.

We estimate the remaining parameters by the SMM approach. In particular, we set the parameter values to minimize the percentage difference with equal weighting between a vector of nine model moments and its data counterpart. Next, we describe each parameter and its most informative moment in detail.[3]

[3] See Table A.4 for their model fit.

Labor disutility shifter (ϕ_w, ϕ_m)**.** We set the labor disutility shifter of production workers ϕ_w to match the average firm size in terms of production workers. In the data, the average firm hires 5.6 production workers. In addition, we choose the labor disutility shifter of managers ϕ_m to match the economy-wide share of managers over total employment of 19%.

Span of control (α)**.** The ratio of production workers to managers partially depends on the span of control parameter. When this parameter increases, the technology converges towards a production function with constant marginal returns to labor, thus decreasing the incentives for managerial delegation. We calibrate this parameter to target that the median multi-layer firm has 3.2 workers per manager.

Efficiency of labor (φ_w, φ_m). We use the efficiency of labor of each occupation to target wages in the data, as these parameters are informative of the marginal product of labor. We use the efficiency of labor of production workers φ_w to match a mean monthly wage of €718 for production workers. Then, we set the efficiency of labor of managers φ_m to match that the wage gap in mean wages between managers and production workers is equal to 0.73 log points.

Dispersion in firm productivity (σ_z)**.** The level of productivity dispersion across firms within a market affects the level of market payroll concentration. We use the standard deviation of firm productivity to target the employment-weighted average HHI across local labor markets of production workers, which equals 0.19.

Within-market substitutability (η_w, η_m)**.** The within-market elasticity parameters are informative of the relationship between firms' productivity and wages for the sub-sample of unconstrained firms in each market. In particular, the inverse labor supply curve in Equation 7 delivers the following equilibrium relationship between (log) wages and (log) employment:

$$\log(w^*_{ijo}) \;=\; \frac{1}{\eta_o}\log(n^*_{ijo}) \;+\; \underbrace{\underbrace{\left(\frac{1}{\theta_o} - \frac{1}{\eta_o}\right)\log\Big(\mathbf{n_{jo}}(n^*_{ijo}, n^*_{-ijo})\Big)}_{\text{Effect of payroll share on wages stems from } \mathbf{n_{jo}}} + \frac{1}{\theta_o}\log(\mathbf{N_o}) \;+\; \log(\mathbf{W_o})}_{\text{Common across firms in a market } jo}.$$

Table 3: Parameterization

Parameter Value	Description	Value	Moment
Panel I: Exogenous Calibration			
γ	Aggregate Frisch elasticity	0.50	Berger et al. (2022)
(κ_w, κ_m)	Household profit shares	(0.81 , 0.19)	Occupational employment share
Panel II: Endogenous Calibration			
$G(m_j)$	Firm distribution		Mean, variance, and mass single-firm
B_{ijw}	Amenities in small markets	6.45	Share workers in markets $M_j \leq 10$
Panel III: SMM Estimation			
A: Preferences			
ϕ_w	Labor disutility shifter: workers	4.03×10^{-5}	Average firm size
ϕ_m	Labor disutility shifter: managers	0.78	Share managers
B: Firm Organization			
α	Span of control	0.76	Median span of control
φ_w	Worker efficiency	3,966	Mean wage of prod. workers
φ_m	Managerial efficiency	2.86	Wage gap managers and prod. workers
σ_z	Std. Dev. firm TFP	0.56	Weighted Mean HHI prod. workers
C: Firm Substitutability			
(θ_w, θ_m)	Across-market firm substitutability	(2.58 , 0.70)	Across-municipality labor supply elasticity
(η_w, η_m)	Within-market firm substitutability	(10.93 , 5.42)	Within-market labor supply elasticity

Sources: The Table reports the quantification of model parameters. Panel I reports the parameters that we calibrate outside the model. Panel II reports the parameters that we calibrate inside the model. Finally, Panel III shows the results from the SMM estimation.

Note that, conditional on common market features, all firms face the same labor supply elasticity η_o for each occupation. This occurs because the effect of the payroll share on the labor supply elasticity shuts down when we control for market employment. We use this insight to obtain a theory-consistent estimate of the across-market elasticity for each occupation. In particular, the previous equation implies the following empirical reduced-form relationship for the inverse labor supply curve:

$$\log(w_{ijo,t}) \;=\; \beta_o \log(n_{ijo,t}) \;+ \mu_{jo,t} \;+\; \nu_{ijo,t}, \tag{17}$$

where $\mu_{jo,t}$ stands for the market-year fixed effects that empirically control for common labor demand and supply shocks across firms in the same market year. Our coefficient of interest is β_o.

In the model, conditional on the sub-sample of unconstrained firms, the OLS regression of Equation 17 separately for each occupation identifies the within-market elasticity as $\eta_o = 1/\hat{\beta}_o$. This regression exploits the cross-sectional variation in employment and wages that uniquely stems from labor demand differences across firms in the same market while keeping their labor supply curve fixed. The intuition is as follows. Firms pay different wages and hire a different number of workers because they are heterogeneous in productivity. Higher productivity, either because firms adopt a multi-layer organization or have a higher TFP, has two equilibrium effects. First, the labor demand curve shifts to the right because the marginal productivity rises. Second, the labor supply curve shifts to the right because the strategic complementarities from Cournot's competition imply that competitors restrict employment. The coefficient β_o absorbs the first effect while market-fixed effects absorb the second effect.

In our baseline specification, we estimate Equation 17 by OLS separately for each occupation. Regarding the sample selection, we restrict the sample to firms paying wages at least one percent higher than the minimum wage in each year. This regression implies a within-market labor supply elasticity of 19.96 for production workers and 6.10 for managers (see Table C.1). In addition, since the error may capture firm-specific labor supply considerations that threat identification, we also provide the results from an IV regression that uses a traditional Bartik instrument in the urban literature to predict firms' employment from national sector employment and initial shares (see Appendix C.5). This alternative specification provides similar results. Overall, we estimate that production workers are three times as responsive to changes in labor demand as managers.

Across-market substitutability (θ_w, θ_m). The across-market firm substitutability parameters govern how greater market productivity translates into higher wages. When firm substitutability is high, total employment in a particular market is highly responsive to increased market productivity. We use an indirect-inference approach for each occupation to match the reduced-form inverse labor supply elasticity from a municipality-level regression. We estimate the following regression:

$$\text{Log}\, w_{m,o,t} = \gamma\, \text{Log}\, L_{m,o,t} + \alpha_{m,o} + e_{m,o,t}, \tag{18}$$

where $w_{m,o,t}$ is the mean wage in municipality m for occupation o in period t, $L_{m,o,t}$ is total employment in that municipality, and $\alpha_{m,o}$ are municipality fixed effects. In the model, the municipality is a collection of local labor markets. We use a municipality-level regression, instead of a market-level regression, to use the standard formulation of the shift-share instrument for employment (Blanchard et al., 1992):

$$\hat{L}_{m,o,t} = \sum_s \Bigg(\underbrace{\frac{L_{i,m,s,o,2007}}{\sum_i L_{i,m,s,o,2007}}}_{\text{Industry-Municipality Share}} \times \underbrace{\sum_i L_{i,s,o,t}}_{\text{National Employment in Sector } s} \Bigg). \tag{19}$$

This instrument predicts employment in each municipality as the exposure-weighted average of national sector employment, where the weights are the employment shares of each sector in the municipality in the initial period. With this approach, we estimate the coefficient γ from within-municipality, across-time variation in wages and employment that arises from national sector employment shocks and the municipality's exposure to them, which plausibly represent exogenous changes in the municipality's labor demand. We also restrict the sample to municipalities whose mean wage is higher than 25 percent of the minimum wage in each respective year. We impose this restriction to exclude municipalities where a high share of firms pay wages close to the minimum wage, as the labor supply elasticity of these firms is not informative of the across-market elasticity because they may pay the minimum wage either before or after the labor demand shock. Table C.2 reports the IV results for each occupation. The implied coefficients are 1.46 for production workers and 0.92 for managers.

To replicate this regression in the model, we first randomly assign markets to each municipality such that we approximate the number of markets that the average municipality has in the data, while keeping a reasonable sample size of municipalities. Then, we simulate two periods, where the second period involves random productivity shocks at the municipality level from a standard log normal distribution with $\sigma = 0.05$. Finally, we choose (θ_w, θ_m) to target the inverse labor supply elasticity γ that results from estimating the regression in Equation (18) with the simulated sample. We infer an across-market elasticity of 1.48 for production workers and 0.93 for managers.

4.2 Discussion on the estimation of labor supply elasticities

Given the importance of the firm-substitutability parameters in estimating wage markdowns, we provide additional evidence on mobility measures that support our findings and compare our estimates with the literature.

The main finding from our estimates is that production workers are significantly more responsive to changes in labor demand than managers. Panel A of Table A.3 shows that production workers have two relevant characteristics, which suggest that they face lower mobility costs and search frictions. This result is consistent with their higher labor supply elasticity. First, young workers are significantly more likely to migrate (Molloy et al., 2011; Kennan and Walker, 2011) and make greater job search efforts (Faberman et al., 2022). Nearly 15 percent of production workers are younger than 25, while only 4 percent of managers are below that age. Second, workers in temporary contracts search harder than workers in permanent contracts (Kahn, 2012). Notably, the share of production workers with temporary contracts is 31 percent, twice that of managers with temporary contracts. In addition, since the labor supply elasticity of production workers is higher, one would expect that they are also more mobile across markets and firms. Panel B of Table A.3 reports that production workers are more likely to change firm, municipality, and sector relative to managers.

Regarding the comparison of our estimates with the literature, we focus on the estimates from two papers that also estimate the firm-substitutability parameters using a general equilibrium model of an oligopsony. Berger et al. (2022) estimate the firm-substitutability parameters using an indirect inference approach to replicate size-dependent labor supply elasticities at the firm level, which they estimate using U.S. microdata and exploiting changes in state corporate taxes. Abstracting from different occupations, they define the local labor market as the combination of a three-digit industry and a commuting zone. They estimate $\theta = 0.42$ and $\eta = 10.85$. Using French data, Azkarate-Askasua and Zerecero (2023) define a local labor market as the intersection between a commuting zone, three-digit industry, and occupation. Nonetheless, they assume the firm-substitutability parameters to be common across occupations. They find an across-market elasticity $\theta = 0.42$ and industry-specific within-market elasticities that range within $\eta \in \{1.22, 4.05\}$. Despite using a different methodology, we find our estimates reasonably close to their results. Our estimates of the within-market elasticities are within the range of their results. Our estimates of the across-market elasticity are somewhat higher, but we find this result consistent with the fact that Portuguese municipalities are smaller than U.S. and French commuting zones.[4]

4.3 Model Fit

Before turning to the counterfactual analysis, we discuss how the model captures untargeted moments of employment, wages, and market concentration. In addition, we exploit large mass-layoff shocks to assess the model's generated employment and wage responses relative to the ones in the data.

Simulating mass-layoff shocks. Mass-layoff shocks involve a sudden, sizable, and enduring reduction in the employment size of a prominent plant within the regional economy. As a result, these events potentially have a significant effect on regional wages and employment due to the initial shock and its spillover effects on the local economy.

Mass-layoff shocks are relevant to assess the performance of employment and wage equilibrium responses in the model, as there is an initial firm-specific labor demand shock that

[4]Comparing mobility rates, nearly 10 percent of workers change municipality in Portugal every year. Instead, the yearly migration rate across U.S. counties, subsets of commuting zones, ranges between 3-6 percent in the same period (Molloy et al., 2011).

affects households' labor supply decisions and other firms' labor demand decisions. Furthermore, the model incorporates two main channels through which areas adjust to this shock: (i) outflows to non-employment and (ii) worker reallocation to other firms, either in the same or a different area. These shocks are helpful for this setting because the empirical evidence suggests that most of the impact of mass-layoff shocks on regional employment depends on the degree of worker reallocation across regions (Foote et al., 2019; Gathmann et al., 2020).[5] Therefore, we use this simulation as a validation exercise of the across- and within-market firm substitutability estimates, which are the key drivers of the allocation of workers across markets.

To identify mass layoffs in the Portuguese data, we use sharp and sizeable drops in the employment of an establishment. More precisely, a drop of 100 workers in any establishment for two consecutive years between 2004 to 2016 is a mass layoff.[6] All mass layoff events are aggregated at the municipal level to construct the treatment, which varies across time, and define the treated municipalities as any experiencing a mass layoff and those who do not in the control group. Next, to quantify the impact on local employment and wages, we use an event-study specification that compares the changes in employment and wages of treated and control municipalities for managers and production workers following a mass layoff. Because the treated municipalities have different characteristics relative to the control ones, we also use covariates in the baseline period that capture the size and structure of the local economy for a cleaner comparison between groups. These variables are the log of regional employment, the share of manufacturing employment, the share of highly educated workers, the share of male workers, and the share of young workers. The main assumption needed in this setup is the parallel trends assumption, stating that in the absence of mass layoffs, treated and

[5]These papers find that geographic mobility accounts for most of the regional employment change. For instance, Gathmann et al. (2020) finds that outflows to non-employment account for only 20 percent of the overall employment decline in German regions following a mass layoff. They also find that geographic mobility protects workers below 50 from suffering employment losses.

[6]We further restrict establishments with more than 1% of local employment in the baseline period, and we do not take into account plant closures to define mass layoffs.

Table 4: Estimates of Mass Layoff Shocks on Municipalities

	Production Workers		**Managers**	
	Model	Data	Model	Data
Panel A: Log Wages				
Event Region	-0.006***	-0.007	-0.005	-0.006
	(0.001)	(0.005)	(0.004)	(0.010)
Panel B: Log Employment				
Event Region	-0.020***	-0.061***	-0.027***	-0.040**
	(0.002)	(0.013)	(0.005)	(0.019)
Observations	200	3,084	200	3,084

Notes: The table reports the empirical (Data) and simulated (Model) results from exploiting mass-layoff shocks. For the Data, we show the average treatment effect of mass-layoff shocks on average municipality wages (Panel A) and total municipality employment (Panel B). The dependent variable is the log municipality's employment and wages. 95% confidence intervals in parentheses.

control groups would evolve similarly; in the Appendix C.6, we show that pre-treatment coefficients are not statistically different from zero, suggesting that this assumption holds.[7]

In the model, we randomly assign (across all the areas) a large productivity shock to firms that fulfill three characteristics that are informative of the firm's importance within the municipality and the size of the shock. First, we restrict the shock to multi-layer firms, as more than 95 percent of plants experiencing mass layoffs hire managers and production workers. Second, we exploit that the average firm carrying out a mass layoff had an average municipality employment share of 4.4 percent one year before the mass layoff. Thus, we condition the random shock on the sub-sample of firms whose municipality employment share falls within a specific range to match the same average size of shocked firms. Third, we choose the magnitude of the productivity shock to target that the average firm undergoing a mass layoff reduces its workforce by nearly 50 percent.

Table 4 reports the empirical and simulated results of the impact of mass-layoff shocks on

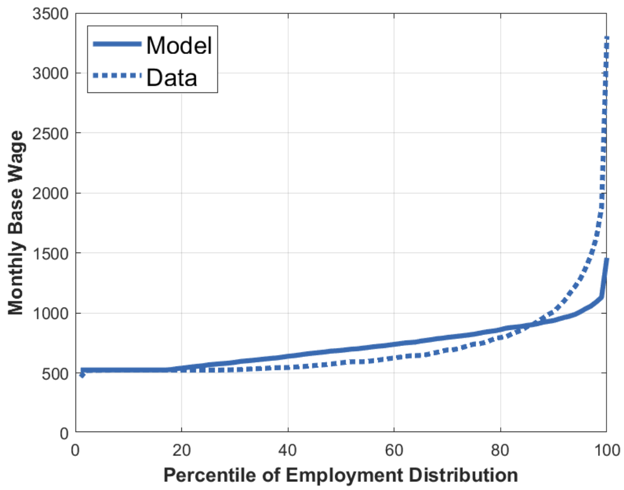
[7] Appendix C.6 explains the empirical specification in more detail.

average municipality wages and total municipality employment. Our empirical results show that municipalities with at least one mass-layoff event experience an average decrease of 6.1 percent in the local employment of production workers and 4 percent in the local employment of managers compared to municipalities without mass layoffs. We do not find a significant impact of mass layoffs on average municipality wages. These results align with Gathmann et al. (2020) for Germany, which finds a drop of 3.7 percent in local employment and an insignificant response in wages following a mass layoff.

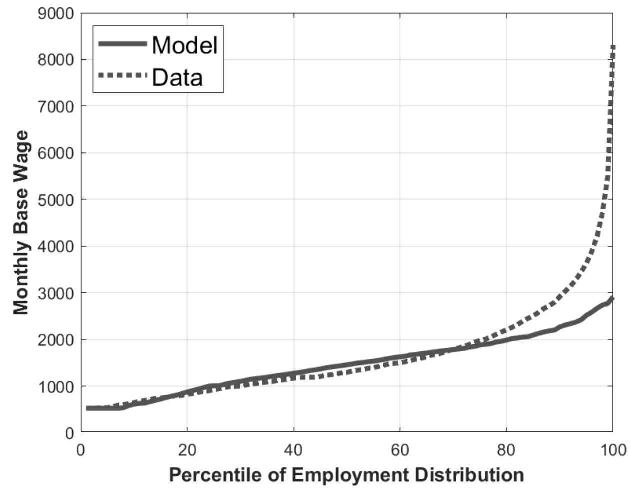
Regarding the results in the simulated data, the model replicates that mass layoffs have a small effect on average municipality wages. The reasons include the downward rigidity in wages due to the presence of minimum wages and the estimated high labor supply elasticity, which imply that most layoff workers relocate to other firms paying higher wages than the shocked firm. In terms of employment, the model accounts for one-third and two-thirds of the decrease in employment of production workers and managers, respectively. Overall, we argue that the model explains a substantial portion of the observed employment decline and attribute the residual to the lack of factors such as agglomeration economies or firm exit. Furthermore, it is worth explaining why the employment response of managers is statistically similar to that of production workers, despite production workers being more mobile. This result stems from the organizational technology and the production function, which incentivizes shocked firms to fire more managers relative to production workers, as firms can adopt a single-layer organization.

Untargeted moments. We start analyzing the wage distribution, which is crucial to quantifying wage dispersion and the repercussions of minimum wage policies. Figure 4 displays the occupation wage distribution in the model and data. The model fits most of the wage distribution for both occupations. Importantly, it replicates the wage distribution at the bottom deciles. This is key for two reasons. First, minimum wage policies have the most substantial employment and wage effects on workers close to the minimum wage. Second, conditional on the occupation, relatively low-paid employees are the most susceptible

Figure 4: Wage Distribution across Occupations



(a) Production workers



(b) Managers

Note: The Graphs show the wage distribution in the model and data for production workers (left) and managers (right).

to monopsony power because they primarily work in small markets. The model slightly underestimates wages in the upper deciles, as we assume that workers are homogeneous in talent within occupations. Nonetheless, high-wage workers primarily work in markets with many firms, where firms have low monopsony power, and their wages are significantly above the minimum wage.

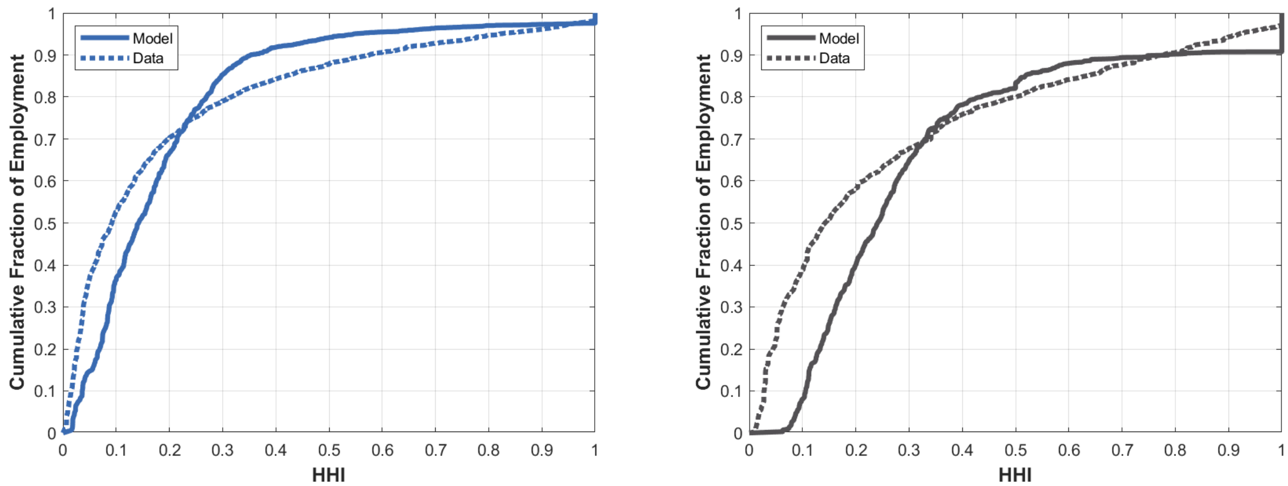
Panel A in Table 5 displays how the model fits moments related to the share of workers whose wages are close to the minimum wage. The model approximates well that about one-fifth of production workers and close to five percent of managers earn the minimum wage. Conditional on being a minimum wage earner, the model also matches that more than 90 percent of employees are production workers, as they are more numerous than managers and more likely to earn the minimum wage. Furthermore, the model captures that most production workers earn less than 700€, a wage close to the binding minimum wage, whereas this proportion falls to almost ten percent for managers. Overall, these facts show that the model can explain that minimum wage policies mostly affect production workers.

Table 5: Untargeted Moments

	Production Workers		**Managers**	
	Model	**Data**	**Model**	**Data**
Panel A: Minimum Wage				
Share minimum wage earners	0.18	0.22	0.08	0.04
Share | Minimum wage earner	0.91	0.96	0.09	0.04
Share wage $\leq$ 700	0.52	0.69	0.14	0.12
Panel B: Firm Size Distribution				
P10	1	1	2	1
P50	2	2	4	1
P90	15	9	11	5
P99	61	59	17	34
Panel C: Firm Organization				
Share Multi-layer	0.26	0.35		
Panel D: Market Concentration				
Weighted mean HHI			0.32	0.27
Weighted mean Max s_{ij}	0.29	0.30	0.42	0.38

Note: The Table reports untargeted moments of the distributions of wages, firm size, and market payroll concentration. For each occupation, we report the statistic from the data and baseline model.

Figure 5: Distribution of Employment across Markets



(a) Production workers

(b) Managers

Note: The Graphs show the cumulative fraction of employment across local labor markets ranked by their level of concentration, i.e. the HHI, in the model and data for production workers (left) and managers (right).

The distribution of firm wages, together with the degree of firm substitutability and local amenities, determines the firm size distribution in the model. This distribution is essential to understand the misallocation of labor across firms that results from high-productivity firms exerting wider markdowns. Panel B in Table 5 shows that the distribution in the model is somewhat similar to that of the data for production workers. Both in the model and data, most firms are small and hire less than two employees. Moreover, the model captures that a handful of firms are relatively large and hire more than 60 workers.

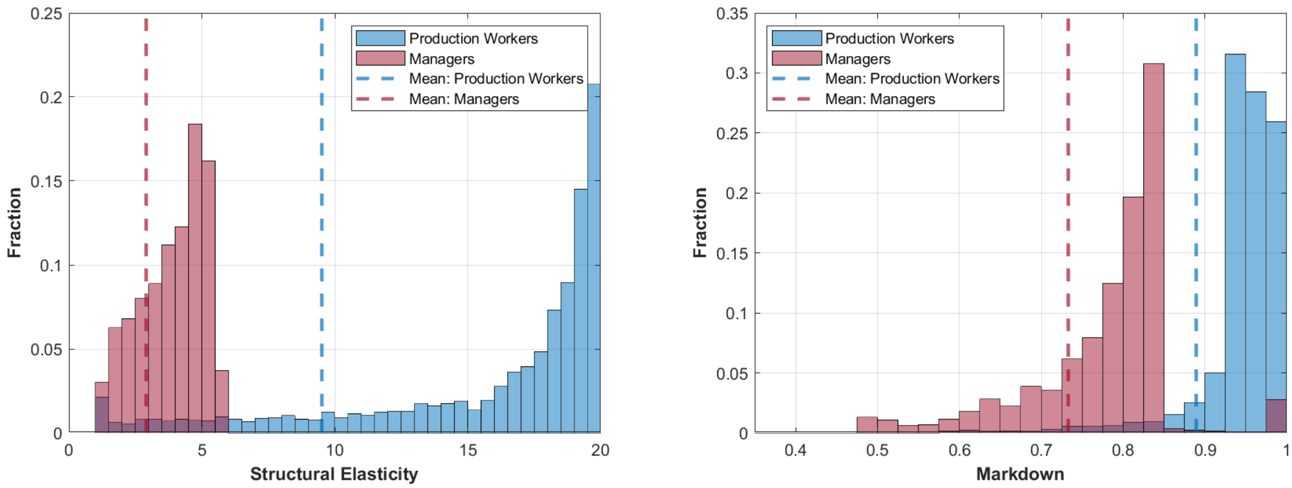
Next, consider the distribution of employment across markets that differ in the level of payroll concentration, which is endogenous in the model due to the labor demand and supply decisions of agents. Figure 5 plots the distribution of employment across markets that differ in terms of the HHI. This distribution is fundamental to measuring the wage markdown of the average worker, as it describes the level of market concentration in which individuals work. The model closely matches both distributions, although we do not target them. Two features stand out. First, the model can explain that most workers sort into a handful of relatively low-concentrated labor markets. These markets attract most workers because they have many firms and pay higher wages, as they have a relatively high degree of market competition. In the case of production workers, these markets are also more attractive because they have relatively higher local amenities. Second, managers are more likely to belong to a market with higher levels of concentration relative to production workers. This heterogeneity occurs because there is a higher proportion of managerial markets with higher levels of concentration, and because managers face higher across-market mobility costs, discouraging them from leaving markets that pay relatively low wages.

Lastly, consider how the model fits two key statistics of market payroll concentration. Panel C in Table 5 shows that the model closely approximates the employment-weighted average of the HHI for managers. Both in the model and data, the average manager works in a market with an HHI of around one-third; the same one would observe with three equally sized firms. In addition to aggregate concentration measures, the model establishes a direct relationship between a wage markdown and the firm's market payroll share. As a result, it is key for the model to predict realistic payroll shares for the biggest firms. The last row in Panel D of Table 5 shows that the model closely fits the average maximum payroll share across markets, where we weigh each market by its employment level. The model captures that the average production worker works for a firm that has nearly one-third of the market payroll, whereas the average manager works for a firm that has about 40 percent of the market payroll.

5 Results

In this section, we quantify the distribution of wage markdowns for both occupations and the effect of monopsony power on output and occupational wage dispersion. We find that firms markdown wages by 11 percent for the average production worker and by 27 percent for the average manager. Relative to the benchmark economy, the competitive allocation leads to an increase of 12 percent in wage dispersion, measured as the ratio of the mean wage

Figure 6: Quantification of Monopsony Power



(a) Structural Labor Supply Elasticity (ε_{ij}) (b) Wage Markdowns (μ_{ij})

Note: The left Graph displays the distribution of the structural labor supply elasticity for production workers (blue) and managers (red). That is the welfare-relevant labor supply elasticity. The right graph plots the distribution of wage markdowns for production workers (blue) and managers (red). The wage markdown is the wedge between the wage and the marginal productivity of labor. Dashed lines display the weighted mean of each variable, where the weight of each firm is its employment size.

of managers to that of production workers, and raises output by 4 percent. The last result stems from an increase of 8 percent in employment and the reallocation of labor towards high-productivity firms, which in the benchmark economy exert the widest wage markdowns.

5.1 Quantification of Monopsony Power

This section shows the distribution of wage markdowns for production workers and managers. In the benchmark economy, wage markdowns are below one due to imperfect firm substitutability, implying that wages are below the marginal revenue product of labor. This wedge represents the efficiency loss from monopsony power, as workers would earn the entire marginal product in the competitive economy.

The left panel in Figure 6 shows the distribution of the structural labor supply elasticity for both occupations. This is the welfare-relevant notion of labor supply elasticity because it determines the wage markdown. We estimate that the employment-weighted mean structural elasticity is equal to 2.92 for managers and 9.51 for production workers. Moreover, the graph shows that most firms face substantially lower labor supply elasticities with respect to managers than to production workers. There are three reasons for this fact. First, both the upper (η_o) and lower bounds (θ_o) of the structural elasticities are lower for managers. That is, we estimate that managers bear greater mobility costs of moving to a firm in another region or industry and greater mobility costs of changing firms within the same region-industry. Second, firms tend to have higher payroll shares of managers than of production workers. Thus, the labor supply elasticity of managers is closer to the across-market elasticity than that of production workers. Put simply, managers find it harder to reallocate toward other firms because more of their alternatives are outside of their current local market. Third, minimum wages constrain low-productivity firms from exercising monopsony power. Since most of these firms only hire production workers, the minimum wage is not an effective policy to reduce wage markdowns for managers, whose wages are significantly further from the minimum wage.

The distribution of the structural elasticities determines that of wage markdowns, which we show in the right panel of Figure 6. Lower structural elasticities translate into lower wage markdowns for managers than for production workers. That is, firms exert greater monopsony power on managerial than on production worker wages. In particular, we estimate an employment-weighted markdown of 11 percent over the wage of the average production worker and a markdown of 27 percent over the wage of the average manager. Therefore, the interaction between monopsony power and firm organization reduces wage dispersion between both occupations. Relative to the efficient economy where salaries equal the marginal product, the implication of this heterogeneity in wage markdowns is that the efficiency and welfare losses from monopsony power mainly stem from distorting managerial employment. Moreover, in the context of policy evaluation, this suggests that additional policies targeting the reduction of monopsony power among low-wage workers, such as an increase in the minimum wage, may not effectively mitigate the welfare losses from monopsony power.

5.2 Welfare and Efficiency Losses from Monopsony Power

In this section, we compute the effect of monopsony power on wage dispersion, as well as the welfare and efficiency losses from monopsony power, to provide a reference for posterior policy evaluation. We compute the efficient allocation by equalizing wages to the marginal product of labor, i.e., wage markdowns equal to one. The middle column in Table 6 summarizes the results from comparing the aggregate outcomes in the efficient relative to the benchmark economy.

Panel A in Table 6 shows the change in employment across occupations. When firms have monopsony power, they internalize that a higher level of employment also involves higher wages, i.e., they internalize an upward-sloping labor supply curve. Then, firms find it optimal to restrict employment relative to the efficient allocation to reduce labor costs. We find that employment would rise by 5 percent in the efficient relative to the benchmark economy. Since firms exert wider markdowns over managerial wages, we estimate that the employment response is twice higher for managers than for production workers. Besides changes in aggregate employment, monopsony power also distorts labor allocation across firms because high-productivity firms restrict employment the most, as they set the widest markdowns. Note that this occurs in our framework both because they face a labor supply elasticity closer to the across-market elasticity and because minimum wages do not constrain their employment policies. Figure A.4 shows the reallocation of employees from less to more productive firms. We observe that firms in the top decile of productivity raise the number of production workers and managers by about 10 and 20 percent, respectively. Thus, monopsony power especially induces a misallocation of managers across firms. Furthermore, we observe the largest employment decline at medium-productivity firms. This phenomenon occurs due to the interplay between firm organization and monopsony power. Namely, the most productive firms need to attract managers to increase their size, who relocate from less productive firms that are nonetheless sufficiently productive to find it optimal to hire managers. Indeed, Figure A.5 shows how a big share of medium-productivity firms remove the managerial layer.

Table 6: Change in Counterfactual Relative to Benchmark

	Efficient Economy	**Minimum Wage Reform**
Panel A: Employment	% Change	% Change
Production workers	4.38	-0.48
Managers	8.72	0.91
Total	5.20	-0.01
Panel B: Firm Organization		
Share multi-layer firms	-23.07	4.11
Median span of control	-4.03	-0.65
Panel C: Wages		
Mean: Production workers	12.18	1.68
Mean: Managers	25.36	0.75
Ratio	11.76	-0.98
Panel D: Efficiency & Welfare		
Output	2.63	0.09
Welfare: Production workers	5.02	-1.23
Welfare: Managers	24.19	1.08

Note: The table reports the percent change in aggregate outcomes in the counterfactual relative to the benchmark economy. In the first column, the counterfactual consists of the efficient economy where wage markdowns to one. In the second column, the counterfactual consists of an economy with a 10 percent higher minimum wage relative to the benchmark economy.

Panel C in Table 6 shows the change in employment-weighted average wages. The increase in labor demand implies that workers in both occupations earn higher salaries. We estimate that the average production worker and manager would earn a 12 percent and 25 percent higher salary, respectively. Therefore, we would observe higher wage dispersion in an efficient economy. Remarkably, the ratio between the mean managerial to production worker wage would rise by 12 percent. Hence, monopsony power is a relevant determinant of wage dispersion between managers and other workers.

Panel D of Table 6 reports the efficiency and welfare losses from monopsony power. We estimate that output would be nearly 3 percent higher in the competitive economy as aggregate employment increases and labor misallocation decreases. The rise in output translates into a much higher welfare gain because, in addition to the output gain, there is a redistribution channel that increases labor income at the expense of profits. Measuring welfare as the consumption equivalent that would make households indifferent between the competitive and benchmark economy, we estimate that welfare increases by 5 and 24 percent for production workers and managers, respectively.

In summary, our findings emphasize the notable impact of monopsony power on efficiency and welfare in the Portuguese economy, despite the government's implementation of a relatively high minimum wage. Regarding the policies that aim to reduce monopsony power, our results point out that the success of these policies relies on their ability to reduce the misallocation of managerial employment across firms. In the following section, we assess the effectiveness of Portuguese minimum wage reforms, which can be thought of as a set of policies that target monopsony power on efficiency and welfare.

5.3 Portuguese Minimum Wage Reforms

Minimum wage policies aim to improve the well-being of low-income workers and reduce the efficiency losses from monopsony power. In this section, we quantify the welfare and efficiency effects of Portuguese reforms that raised the real minimum wage by 10 percent between 2016 and 2019.[8] We focus on this period because it involves a meaningful and permanent minimum wage increase, and it coincides right after our calibration period and before the COVID-19 crisis.

Our model captures four channels through which minimum wages alter wages and employment. First, minimum wages directly change employment and wages at constrained firms. Consistent with empirical evidence that finds both positive and negative effects (see, for instance, Chapter 4 in Dube, 2019), the minimum wage has a positive impact on employment when it remains below a certain threshold, and it has a negative effect after exceeding it. Second, there are also spillover effects on unconstrained firms, which increase wages because their competitors pay higher wages (e.g., Engbom and Moser, 2022; Giupponi et al., 2022). Third, there is a reallocation effect because the rise in the minimum wage leads to employment losses at unproductive firms, part of which relocates to more productive firms (e.g., Dustmann et al., 2022). The channels mentioned above are also included in Berger et al. (2023b). Moreover, our model includes a novel fourth channel, which implies that the minimum wage has spillover effects on managerial employment, despite these workers earning wages significantly above the minimum wage, due to its influence on the organization of work of firms. In particular, the rise in the minimum wage incentivizes multi-layer firms to decrease the span of control of managers because production workers become more expensive relative to managers. Furthermore, it incentives sufficiently productive firms to add a managerial layer for the same reason, alongside reallocating production workers from less productive firms.

The rightmost column in Table 6 reports the percent change in aggregate outcomes in the minimum wage counterfactual relative to the benchmark economy. Panel A in Table 6 shows that the increase in the minimum wage has a minor effect on overall employment, but it affects production workers and managers differently. On the one hand, the employment

[8] In nominal terms, the Portuguese minimum wage increased from 530€ to 600€. Adjusting for the CPI, where we set the base year to 2010, it implies an increase of 10 percent from 525€ to 578€. See https://www.dgert.gov.pt/evolucao-da-remuneracao-minima-mensal-garantida-rmmg.

level of production workers falls by around 0.5 percent. The reason is that most low-wage production workers work at low-productivity firms, which exerts narrow markdowns. For instance, in the baseline economy, the average markdown over the wage of production workers earning less than the counterfactual minimum wage is smaller than 5 percent. On the other hand, the employment level of managers rises by almost 1 percent. The demand for managers increases because production workers become more expensive relative to managers and because of the employment reallocation towards high-productivity firms. To understand this result, we show in Panel B of Table 6 how two moments of firm organization change after the minimum wage reform. First, the share of multi-layer firms increases by 4 percent, mainly because high-productivity firms have more incentives to hire managers due to the reallocation of production workers (see Figure A.6). Second, the median span of control decreases by 0.65 percent. Hence, our results show that considering firms' managerial delegation decisions, minimum wages have heterogeneous effects across occupations.

Next, consider the impact of the minimum wage reform on wages, which we show in Panel C. We estimate an increase of 1.68 percent in the average wage of production workers. Moreover, the increase in labor demand for managers also implies an increase of almost 1 percent in the average wage of managers. Since this increase is lower than that of production workers, we find that minimum wages effectively reduce the managerial wage gap.

Lastly, consider the effect of the minimum wage reform on efficiency and welfare in Panel D. Despite the fall in overall employment, output slightly increases by 0.09 percent. The reason is that labor productivity increases because more firms adopt a multi-layer organization, which is more productive. In addition, employment rises at high-productivity firms, further contributing to the increase in productivity (see Figure A.7). Regarding the change in welfare, managers get a welfare gain of about one percent from the reform, which entirely arises from the increase in consumption. On the contrary, the minimum wage reform has a negative effect of nearly 1.25 percent on the welfare of production workers. Despite a slight increase in consumption, welfare decreases because the reform induces a reallocation of labor across firms that significantly rises the disutility of the labor supply. For instance, we interpret this last finding as a situation wherein the reallocation of labor across firms involves substantial search costs, increases commuting costs, and diminishes the idiosyncratic taste of production workers toward firms, all in an effort for workers to remain employed.

To sum up, the minimum wage reform has small effects on aggregate employment and output, and it fails to address the welfare losses from monopsony power. Remarkably, firm organization sheds light on the underlying mechanisms driving these results and the heterogeneous effect of the reform across occupations. In terms of efficiency, the minimum wage contributes to a better allocation of labor across firms, especially managers. Moreover, it enhances managerial employment. However, the decrease in the employment level of production workers mostly offsets the aforementioned gains. In terms of welfare, the reform makes production workers worse off, even though they presumably constitute the main target of the policy. Two factors contribute to this finding: The combination of a high baseline minimum wage, which is already binding on about one out of five production workers, as well as the limited monopsony power of unconstrained firms employing low-wage workers. In contrast, managers experience a welfare gain, as most of them earn salaries above the minimum wage and benefit from the spillover effects of the reform.

6 Conclusion

The paper's main result shows that markdown heterogeneity and the organization of work within firms generate wage dispersion between managers and production workers, leading to economic implications in terms of wage dispersion, misallocation, and the effect of minimum wage policies. We find that firms exert wider monopsony power over managers than production workers. Thus, monopsony power mainly distorts the level and allocation of managerial employment across firms. As a result, the managerial wage gap would increase by 12 percent in an efficient economy where firms do not exert monopsony power. The efficient economy provides substantial welfare gains for both occupations, which stem from efficiency gains and redistribution of firm profits to labor income. Lastly, we study whether minimum wage policies address the welfare losses from monopsony power. We find that recent Portuguese reforms that increase the minimum wage by 10 percent bring about welfare losses for production workers, as a big share of these work at low-productivity firms that exert narrow markdowns. On the contrary, the reforms benefit managers whose wages and employment levels increase due to a labor reallocation effect.

We consider two valuable extensions for future research. Empirically, we think that market concentration and mobility costs may also show a systematic relationship with other important characteristics of the market. For instance, the skill level required in the job, the age of workers, or whether the market is formal or not in developing countries. Theoretically, we believe that future model extensions should consider worker flows across occupations and occupation-specific minimum wages.

References

Ahlfeldt, G. M., Redding, S. J., Sturm, D. M., and Wolf, N. (2015). The economics of density: Evidence from the berlin wall. *Econometrica*, 83(6):2127–2189.

Ahlfeldt, G. M., Roth, D., and Seidel, T. (2022). Optimal minimum wages.

Autor, D. H. (2014). Skills, education, and the rise of earnings inequality among the "other 99 percent". *Science*, 344(6186):843–851.

Azar, J., Marinescu, I., and Steinbaum, M. (2022). Labor market concentration. *Journal of Human Resources*, 57(S):S167–S199.

Azar, J., Marinescu, I., Steinbaum, M., and Taska, B. (2020). Concentration in us labor markets: Evidence from online vacancy data. *Labour Economics*, 66:101886.

Azkarate-Askasua, M. and Zerecero, M. (2023). Union and firm labor market power. *Working Paper*.

Bamford, I. (2021). Monopsony power, spatial equilibrium, and minimum wages. *Job Market Paper*.

Bartelsman, E., Haltiwanger, J., and Scarpetta, S. (2013). Cross-country differences in productivity: The role of allocation and selection. *American economic review*, 103(1):305–334.

Bartik, T. J. (1991). *Who benefits from state and local economic development policies? Kalamazoo, MI: WE Upjohn Institute for Employment Research*.

Bassanini, A., Bovini, G., Caroli, E., Casanova Ferrando, J., Cingano, F., Falco, P., Felgueroso, F., Jansen, M., Martins, P. S., Melo, A., et al. (2023). Labour market concentration, wages and job security in europe. *Nova SBE Working Paper Series*, (654).

Benmelech, E., Bergman, N. K., and Kim, H. (2020). Strong employers and weak employees: How does employer concentration affect wages? *Journal of Human Resources*, pages 0119–10007R1.

Berger, D., Herkenhoff, K., and Mongey, S. (2022). Labor market power. *American Economic Review*, 112(4):1147–93.

Berger, D. W., Herkenhoff, K. F., Kostøl, A. R., and Mongey, S. (2023a). An anatomy of monopsony: Search frictions, amenities and bargaining in concentrated markets. Technical report, National Bureau of Economic Research.

Berger, D. W., Herkenhoff, K. F., and Mongey, S. (2023b). Minimum wages, efficiency and welfare. Technical report.

Bhaskar, V., Manning, A., and To, T. (2002). Oligopsony and monopsonistic competition in labor markets. *Journal of Economic Perspectives*, 16(2):155–174.

Blanchard, O. J., Katz, L. F., Hall, R. E., and Eichengreen, B. (1992). Regional evolutions. *Brookings papers on economic activity*, 1992(1):1–75.

Borusyak, K., Hull, P., and Jaravel, X. (2022). Quasi-experimental shift-share research designs. *The Review of Economic Studies*, 89(1):181–213.

Caliendo, L., Mion, G., Opromolla, L. D., and Rossi-Hansberg, E. (2020). Productivity and organization in portuguese firms. *Journal of Political Economy*, 128(11):4211–4257.

Caliendo, L. and Rossi-Hansberg, E. (2012). The impact of trade on organization and productivity. *The quarterly journal of economics*, 127(3):1393–1467.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

Chetverikov, D., Larsen, B., and Palmer, C. (2016). Iv quantile regression for group-level treatments, with an application to the distributional effects of trade. *Econometrica*, 84(2):809–833.

David, H., Katz, L. F., and Kearney, M. S. (2006). The polarization of the us labor market. *American economic review*, 96(2):189–194.

Davis, S. J., Haltiwanger, J., Handley, K., Jarmin, R., Lerner, J., and Miranda, J. (2014). Private equity, jobs, and productivity. *American Economic Review*, 104(12):3956–3990.

Dodini, S., Lovenheim, M. F., Salvanes, K. G., and Willén, A. (2023). Monopsony, job tasks, and labor market concentration. Technical report, National Bureau of Economic Research.

Drechsel-Grau, M. (2023). Employment and reallocation effects of higher minimum wages. *Working Paper*.

Dube, A. (2019). Impacts of minimum wages: review of the international evidence. *Independent Report. UK Government Publication*, pages 268–304.

Dustmann, C., Lindner, A., Schönberg, U., Umkehrer, M., and Vom Berge, P. (2022). Reallocation effects of the minimum wage. *The Quarterly Journal of Economics*, 137(1):267–328.

Engbom, N. and Moser, C. (2022). Earnings inequality and the minimum wage: Evidence from brazil. *American Economic Review*, 112(12):3803–47.

Faberman, R. J., Mueller, A. I., Şahin, A., and Topa, G. (2022). Job search behavior among the employed and non-employed. *Econometrica*, 90(4):1743–1779.

Foote, A., Grosz, M., and Stevens, A. (2019). Locate your nearest exit: Mass layoffs and local labor market response. *ILR Review*, 72(1):101–126.

Garcia-Santana, M. and Pijoan-Mas, J. (2014). The reservation laws in india and the misallocation of production factors. *Journal of monetary economics*, 66:193–209.

Garicano, L. (2000). Hierarchies and the organization of knowledge in production. *Journal of political economy*, 108(5):874–904.

Garicano, L., Lelarge, C., and Van Reenen, J. (2016). Firm size distortions and the productivity distribution: Evidence from france. *American Economic Review*, 106(11):3439–79.

Garicano, L. and Rossi-Hansberg, E. (2006). Organization and inequality in a knowledge economy. *The Quarterly journal of economics*, 121(4):1383–1435.

Gathmann, C., Helm, I., and Schönberg, U. (2020). Spillover effects of mass layoffs. *Journal of the European Economic Association*, 18(1):427–468.

Giupponi, G., Joyce, R., Lindner, A., Waters, T., Wernham, T., and Xu, X. (2022). The employment and distributional impacts of nationwide minimum wage changes. *Mimeo*.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624.

Grobovsek, J. (2020). Managerial delegation, law enforcement, and aggregate productivity. *The Review of Economic Studies*, 87(5):2256–2289.

Heise, S. and Porzio, T. (2023). Labor misallocation across firms and regions. Technical report.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and manufacturing tfp in china and india. *The Quarterly journal of economics*, 124(4):1403–1448.

Hurst, E., Kehoe, P. J., Pastorino, E., and Winberry, T. (2022). The distributional impact of the minimum wage in the short and long run. Technical report, National Bureau of Economic Research.

Jarosch, G., Nimczik, J. S., and Sorkin, I. (2023). Granular search, market structure, and wages. *Working Paper*.

Kahn, L. M. (2012). Temporary jobs and job search effort in europe. *Labour Economics*, 19(1):113–128.

Kambourov, G. and Manovskii, I. (2009). Occupational specificity of human capital. *International Economic Review*, 50(1):63–115.

Karabarbounis, L., Lise, J., and Nath, A. (2022). Minimum wages and labor markets in the twin cities. Technical report, National Bureau of Economic Research.

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

Kennan, J. and Walker, J. R. (2011). The effect of expected income on individual migration decisions. *Econometrica*, 79(1):211–251.

Lucas Jr, R. E. (1978). On the size distribution of business firms. *The Bell Journal of Economics*, pages 508–523.

MacKenzie, G. (2021). Trade and market power in product and labor markets. Technical report, Bank of Canada Staff Working Paper.

Manning, A. (2013). Monopsony in motion. In *Monopsony in Motion*. Princeton University Press.

Mariscal, A. (2020). Firm organization and information technology: Micro and macro implications. *Working Paper*.

Martins, P. S. (2018). Making their own weather? estimating employer labour-market power and its wage effects. *CGR Working Paper 95*.

Molloy, R., Smith, C. L., and Wozniak, A. (2011). Internal migration in the united states. *Journal of Economic perspectives*, 25(3):173–196.

Monte, F., Redding, S. J., and Rossi-Hansberg, E. (2018). Commuting, migration, and local employment elasticities. *American Economic Review*, 108(12):3855–90.

Neal, D. (1995). Industry-specific human capital: Evidence from displaced workers. *Journal of labor Economics*, 13(4):653–677.

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

Santamaria, C. (2023). Small teams in big cities: Are technology differences across cities optimal? *Working Paper*.

Sokolova, A. and Sorensen, T. (2021). Monopsony in labor markets: A meta-analysis. *ILR Review*, 74(1):27–55.

Sullivan, P. (2010). Empirical evidence on occupation and industry specific human capital. *Labour economics*, 17(3):567–580.

Tamkoç, N. M. (2022). Production complexity, talent misallocation and development. *Working Paper*.

Online Appendix

A Additional Figures and Tables

Table A.1: Classification of Occupations

Level	Tasks	Skills
Top Management	Definition of the firm general policy or consulting on the organization of the firm; strategic planning; creation or adaptation of technical, scientific and administrative methods or processes	Knowledge of management and coordination of firms fundamental activities; knowledge of management and coordination of the fundamental activities in the field to which the individual is assigned and that requires the study and research of high responsibility and technical level problems
Middle Management	Organization and adaptation of the guidelines established by the superiors and directly linked with the executive work	Technical and professional qualifications directed to executive, research, and management work
Supervisors	Orientation of teams, as directed by the superiors, but requiring the knowledge of action processes	Complete professional qualification with a specialization
Higher-skilled Profesisonals	Tasks requiring a high technical value and defined in general terms by the superiors	Complete professional qualification with a specialization adding to theoretical and applied knowledge
Skilled Professionals	Complex or delicate tasks, usually not repetitive, and defined by the superiors	Complete professional qualification implying theoretical and applied knowledge
Semi-skilled Professionals	Well defined tasks, mainly manual or mechanical (no intellectual work) with low complexity, usually routine and sometimes repetitive	Professional qualification in a limited field or practical and elementary professional knowledge
Non-skilled Professionals	Simple tasks and totally determined	Practical knowledge and easily acquired in a short time

Sources: (i) *Decreto-Lei nº. 121/78 de 2 de Junho, Ministério do Trabalho*, (ii) Caliendo et al. (2020).

Table A.3: Mobility and Sample Characteristics

	(1) Production Workers	(2) Managers
	Mean	Mean
Share Age≤25	0.13	0.04
Share Age≤30	0.28	0.19
Share Temporary	0.31	0.16
Share Change Firm	0.24	0.20
Share Change Municipality	0.11	0.10
Share Change Sector	0.20	0.17
Observations	24,362,091	5,128,131

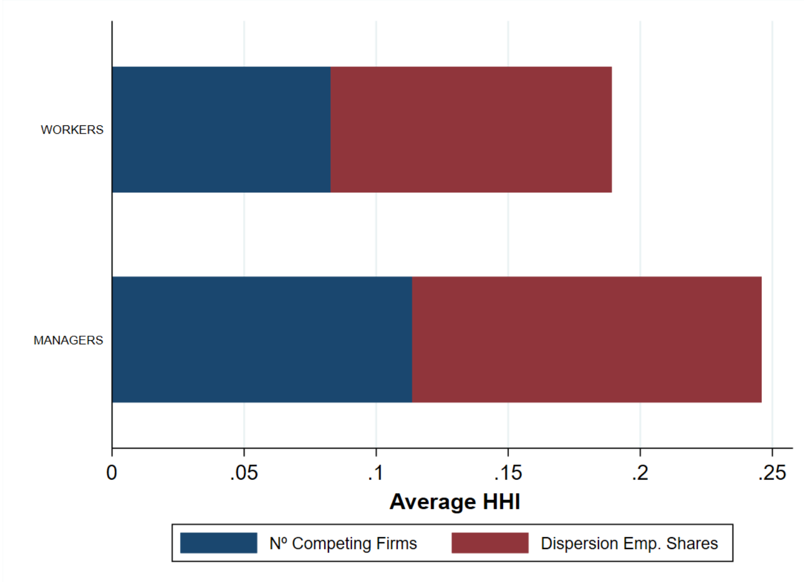
Source: Elaboration based on QP.

Table A.2: Classification of Occupations

Level	**Share (%)**	**Share Hierarchy (%)**	**Mean Wage**	**Std. dev. Wage**
Managers	19.19	100	1,850	1,398
Top Management	7.97	41.55	2,226	1,778
Middle Management	5.96	31.08	1,658	1,033
Supervisors and Team Leaders	5.25	27.37	1,498	862
Workers	80.81	100	808	842
Higher-skilled Professionals	8.07	9.98	1,367	2,318
Skilled Professionals	40.44	50.04	841	443
Semi-skilled Professionals	21.48	26.58	661	281
Non-skilled Professionals	10.83	13.40	565	241

Source: elaboration based on Quadros de Pessoal.

Figure A.1: Decomposition Average HHI



Source: elaboration based on QP.

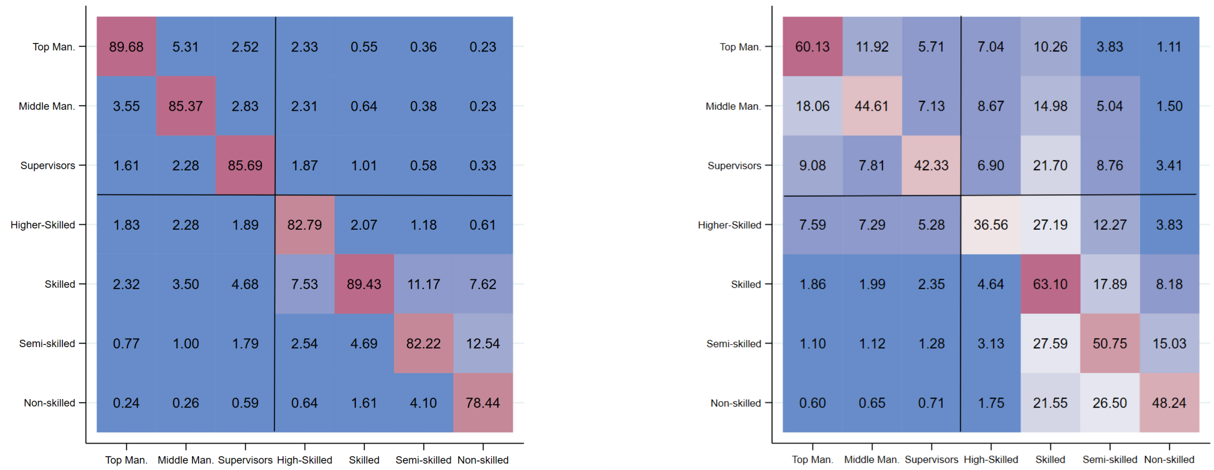
Note:

Table A.4: Targeted Moments

Parameter Value	Description	Moment	Model	Data
A: Preferences				
ϕ_m	Labor disutility shifter: production workers	Average firm size	5.67	5.63
ϕ_w	Labor disutility shifter: managers	Ratio managers to workers	0.26	0.24
B: Production Organization				
α	Span of control	Median span of control	3.41	3.26
F	Fixed cost multi-layer	Share multi-layer	0.33	0.34
φ_w	Worker efficiency	Mean wage of workers (€)	734	718
φ_m	Managerial efficiency	Wage gap managers and workers	0.74	0.73
σ_z	Std. Dev. firm TFP	Weighted mean HHI workers	0.19	0.19
C: Firm Substitutability				
$(\theta_w\,,\theta_m)$	Across-market firm substitutability	Across-municipality LS elasticity	(2.17 , 1.00)	(2.31 , 1.00)
$(\eta_w\,,\eta_m)$	Within-market firm substitutability	Within-market LS elasticity	(10.92 , 5.42)	(10.92 , 5.42)
D: Firm Distribution				
Mass $m_j = 1$	Share single-firm markets	Mass single-firm markets	0.29	0.29
ζ_0	Scale Pareto distribution	Mean Nº firms	17.87	17.63
ζ_1	Shape Pareto distribution	Std. Dev. Nº firms	72.65	68.25

Note: The Table reports the vector of parameters estimated using the SMM approach and the calibrated firm distribution with their respective moment description and fit.

Figure A.2: Transition Probabilities

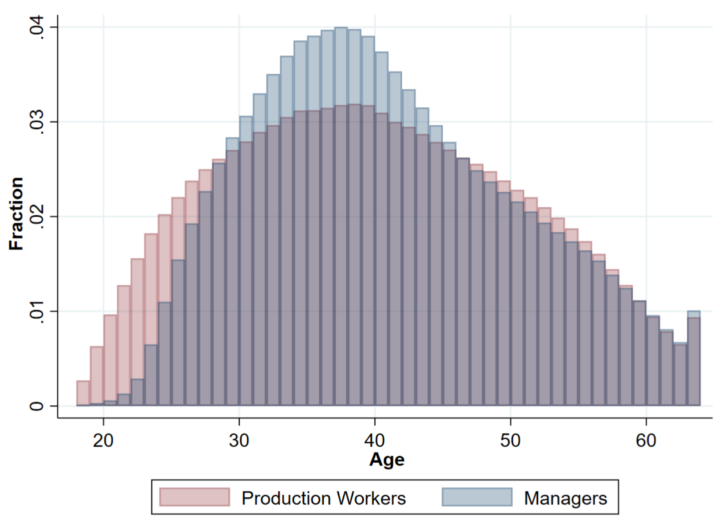


(a) Unconditional

(b) Conditional

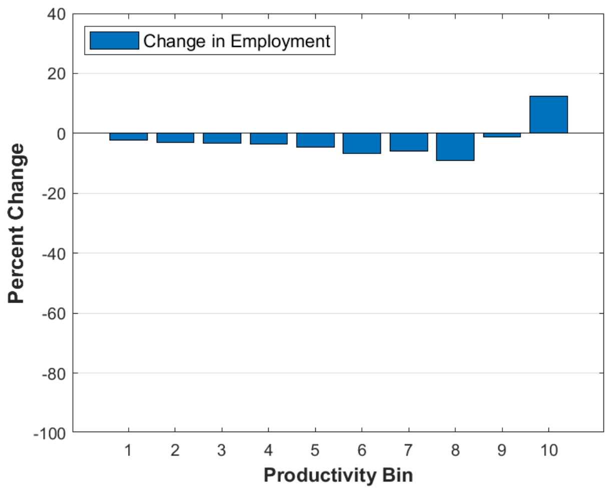
Note: The Graphs displays the transition probabilities of changing sub-occupation. The left panel shows the unconditional transition probability, whereas the right panel shows the transition probability conditional on changing firm. The black lines delimit the quadrants of moving across or within the two broad occupation categories (managers and production workers), where the top left and right bottom quadrants represent within-occupation transitions. The vertical axis represents the sub-occupation before the transition, and the horizontal axis the sub-occupation afterwards.

Figure A.3: Age Distribution across Occupations

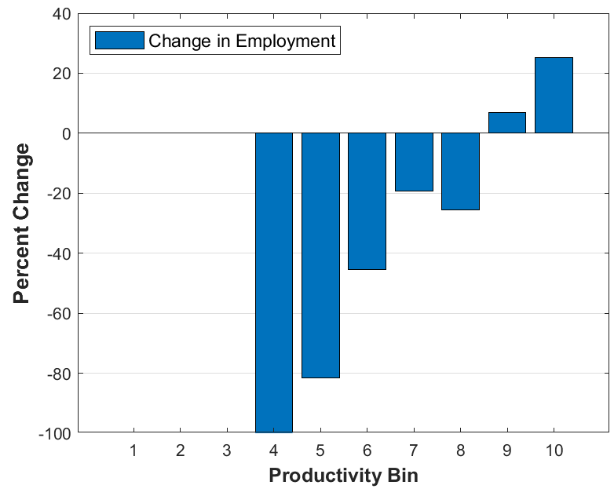


Note: The Graph displays the age distribution across occupations.

Figure A.4: Employment Reallocation in the Efficient Economy



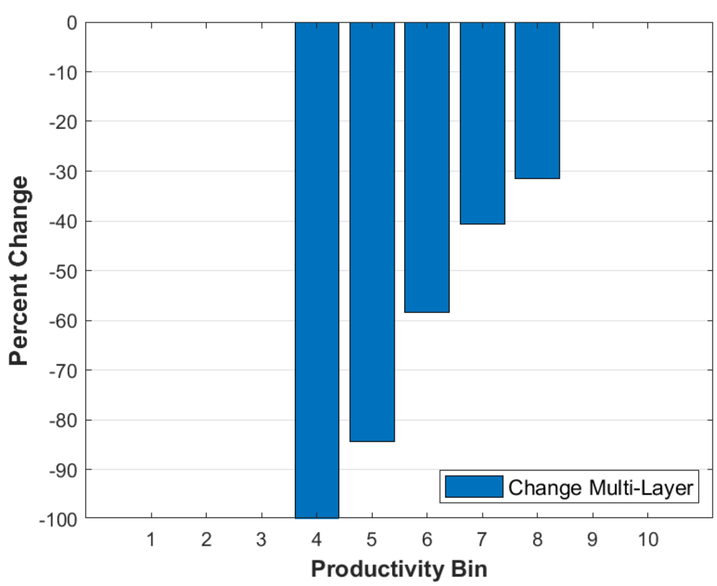
(a) Production Workers



(b) Managers

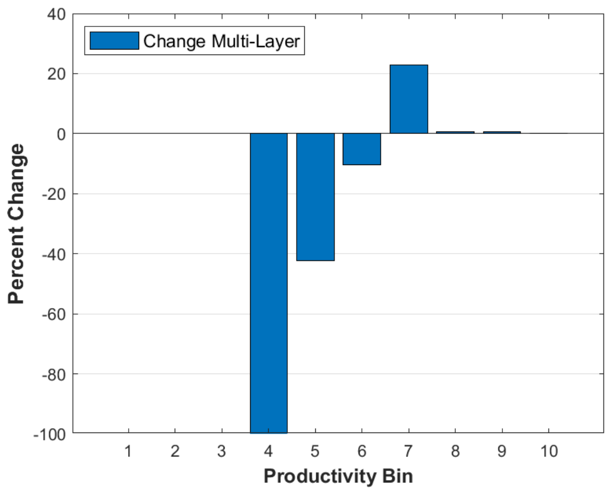
Note: The Graph plots the percent change in employment of production workers (left) and managers (right) across firms in the efficient relative to the benchmark economy. We classify firms into ten bins according to their productivity, where a higher bin implies higher productivity.

Figure A.5: Change in the Share of Multi-layer Firms in the Efficient Economy

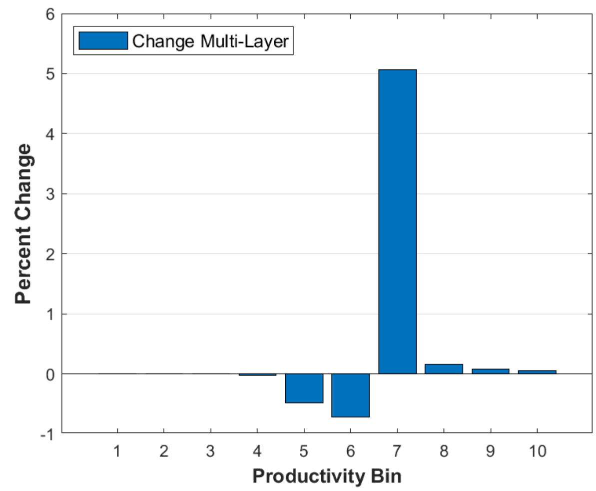


Note: Comparing the efficient economy to the benchmark economy, the Graph plots the percent change in the share of multi-layer firms across firms. We classify firms into ten bins according to their productivity, where a higher bin implies higher productivity.

Figure A.6: Effect of Minimum Wage Reform on the Share of Multi-layer Firms



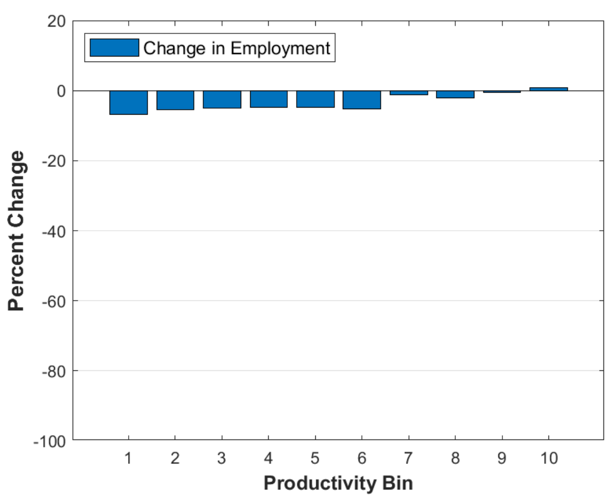
(a) Over Total Firms in Productivity Bin



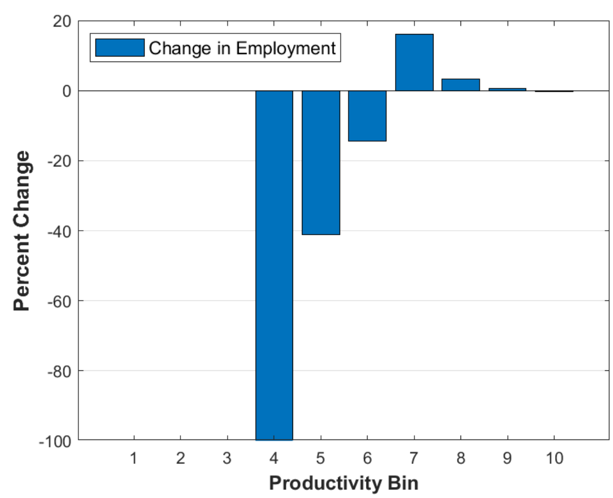
(b) Over Total Firms

Note: The Graph plots the percent change in the share of multi-layer firms, where we compare the change over the total number of firms (right) and over the total number of firms in the same productivity bin (left).

Figure A.7: Effect of Minimum Wage Reform on Employment Reallocation



(a) Production Workers



(b) Managers

Note: The Graph plots the percent change in employment of production workers (left) and managers (right) across firms in the competitive relative to the benchmark economy. We classify firms into ten bins according to their productivity, where a higher bin implies higher productivity.

B Data

This section provides a detailed description of the data, the market definition, and the methodology to measure market concentration using the Herfindahl-Hirschman Index (HHI).

B.1 Quadros de Pessoal

Our primary data source is *Quadros de Pessoal* (QP), an annual census of private sector employees conducted by the Portuguese Ministry of Employment. This census provides matched employer-employee data on all firms based in Portugal with at least one worker.[9] The database incorporates unique time-invariant identifiers for each firm, establishment, and worker entering the report, which allows tracking them over time. In particular, we cover the period from 2010 to 2016.

The worker-level data contains information on each firm's employees as of the October reference week. The variables include age, occupation, monthly earnings, and hours worked. At the firm level, we have information on the industry, the headquarters location, and all its establishments.

Regarding the sample selection, we exclude workers younger than 18 or older than 64, those working outside of continental Portugal, and those working in agriculture, forestry, fishing, or mining industries. We also exclude apprentices, workers with missing information on earnings or occupation, and workers with misreported identifiers.[10] Finally, we drop chief executive officers because their market is not local, which is a core feature of the theory in this paper. This selection results in 3,492,148 workers and 13,036,426 worker-year observations.

[9] The census excludes the public administration and non-market services.

[10] Most workers with missing earnings include unpaid family members and owners of the firm. In addition, workers with misreported identifiers (e.g., duplicated) account for about 2% of the sample.

B.2 Market Definition

We classify labor markets based on three observable characteristics of the job: geography, industry, and occupation. This classification stems from the fact that workers are more attached to their current labor market because of imperfect substitutability of skills across jobs and sectors and imperfect geographical mobility (Neal, 1995; Kambourov and Manovskii, 2009; Sullivan, 2010; Monte et al., 2018). In particular, we define two broad occupations, i.e., managers and production workers, and define a local labor market for each occupation as the intersection of the geography (Municipality) and industry (2-digit NACE). This selection results in 13,890 and 11,765 local labor markets for workers and managers, respectively.

The geographical dimension reflects that workers likely stay in their current labor market due to geographical mobility costs. We use the municipality or *concelho* administrative division as the benchmark geographic unit, which splits the country into 278 areas of an average of 320 square kilometers. Within each geography, workers are more likely to stay in firms of the same sector due to the imperfect substitutability of human capital across sectors. We use the 2-digit NACE classification of industries as a baseline measure. This includes 78 different economic sectors such as "*Manufacture of food products*" or "*Accommodation and food service activities*". Given that our model does not distinguish between across-industry and across-region mobility, we use these baseline definitions because worker transitions are similar in both cases. In particular, the unconditional across-municipality and across-industry annual transition probabilities are both 9.8 percent.

Regarding the occupational definition, the Portuguese law obliges firms to assign their workers to an occupational category based on tasks performed and skills required so that each category considers the level of the worker within the firm's hierarchy in terms of increasing responsibility and task complexity. We follow a hierarchical classification similar to (Caliendo et al., 2020). In particular, we partition professional categories into two layers. We assign "top executives", "intermediary executives", and "supervisors, team leaders to the management layer". In addition, we assign "higher-skilled professionals", "skilled profes-

Table A.5: Summary Statistics at the Establishment Level

	Mean	P10	P25	P50	P75	P90
Production Workers						
Monthly Wage	718	518	588	756	1,082	2,159
Managers						
Monthly Wage	1,698	937	1,346	2,059	3,065	6,441
Span of Control	8	1	3	8	17	70

Source: Elaboration based on Quadros de Pessoal.

Note: The Table reports the mean, 10th, 25th, 50th, 75th, and 90th percentile of the individual distribution of wages for managers and non-managers. Wages are base wages (excluding supplementary payments) expressed in full-time equivalent units. In addition, the table reports the same distributional moments for the span of control, which we define as the ratio of non-managers to managers within an establishment.

sionals", "skilled professionals" "semi-skilled professionals", and "non-skilled professionals" to the bottom layer. To distinguish between managers and other occupations, the critical difference is that managers are responsible for the organizational policies of the firm and their adaptation, which require a high degree of qualification in terms of direction, guidance, and coordination of the firm fundamental activities.[11]

B.3 Summary Statistics

Our classification of occupations implies that 19 percent of workers are managers, while the remaining 81 percent are production workers. Table A.5 reports summary statistics of the wage distribution for each occupation. Along the distribution, managers earn higher wages than production workers, and this gap particularly widens for high-paid workers. In the bottom quartile, managers earn around twice as much as production workers. In the top quartile, managers earn nearly three times as much as production workers.

This wage gap arises even though about two-thirds of managers are not top executives but supervisors, team leaders, or intermediary executives (see Table A.2). We measure the

[11] See Table A.1 for further information about the categories of the occupational classification, which is based on *Decreto-Lei nº. 121/78 de 2 de Junho, Ministério do Trabalho.*

number of workers a manager has under his charge, i.e., the span of control, with the ratio of non-managers to managers in each establishment. In half of the establishments, managers have a span of control lower than eight workers, and only one-fourth of establishments have managers with a span of control greater than seventeen workers. These results highlight that most establishments assign to their managers a small span of control.

To sum up, we find substantial wage dispersion between managers and production workers. The literature on income inequality mainly attributes wage differences between groups to productivity-enhancing forces such as skill-biased technologies (Katz and Murphy, 1992; Autor, 2014), task-biased technologies (David et al., 2006), and trade specialization (Chetverikov et al., 2016). In the next section, we show evidence of an additional potential force behind both types of wage inequality: heterogeneity in market payroll concentration as a proxy for monopsony power.

B.4 Measuring Market Payroll Concentration

Our baseline measure of market concentration is the HHI. Given the employment n_{ij} and wage w_{ij} level at firm i in a local labor market j, we define the HHI in the market as:

$$\text{HHI}_j := \sum_{i=1}^{N_j} s_{ij}^2 = \frac{1}{N_j} + \sum_{i=1}^{N_j} \left(s_{ij} - \frac{1}{N_j} \right)^2, \tag{20}$$

$$s_{ij} := \frac{w_{ij} n_{ij}}{\sum_{i=1}^{N_j} w_{ij} n_{ij}}. \tag{21}$$

Where N_j is the number of establishments in market j and s_{ij} stands for the payroll share of the firm i. The HHI equals the average payroll market share weighted by the payroll share itself. The index ranges from $\frac{1}{N}$ to 1, where a low value reflects low concentration or many firms having similar payroll shares. Note that this index gives more weight to larger establishments, especially penalizing markets where a few firms have a large share of the market payroll. The rightmost equality of Equation (20) shows that the HHI has an economically meaningful decomposition into two concentration sources. The left term reflects that the concentration level decreases with the *number of establishments*, given that they have identical payroll shares. Given the number of firms, the right term shows that the index increases when there is a higher *variance in payroll shares* across establishments relative to the symmetric case.

B.5 Stylized Facts

Fact #1: Managers tend to work in markets with higher payroll concentration. Figure 1 plots the estimated kernel density of the market level HHI for each occupation and their respective cumulative share of employment. We observe that managers are more likely to belong to a relatively highly concentrated market than production workers. This occurs because their markets are more likely to present higher payroll concentration (see Figure 1a) and because there is a higher share of managers working there (see Figure 1b).

Table 1 reports particular moments of the distribution of the HHI. The employment-weighted average of the HHI is 38 percent higher for managers than for production workers. In particular, the average HHI is 0.27 for managers and 0.20 for production workers. To provide context for these numbers, one would observe the same level of concentration with four and five equally-sized establishments, respectively. Table 1 also reports moments from the distribution of the median and maximum payroll share across local labor markets. The average largest establishment in a local labor market accumulates about 8 percentage points more market payroll in the market of managers than in the market of production workers. Besides, in most labor markets, the median establishment has a payroll share in managerial markets which is twice as large as the payroll share in the markets of production workers.

The aforementioned facts show that firms' employment policies lead to greater market payroll concentration in managerial occupations. To understand the sources that bring about this observation, we next look into facts related to the organization of work of firms.

Fact #2: Nearly two times as many firms compete for production workers as

Table A.6: Distributional Moments of Establishments across Local Markets

	Mean	P10	P25	P50	P75	P90
Managers						
Nº Establishments	116	3	10	38	153	379
Production Workers						
Nº Establishments	221	5	16	63	236	542

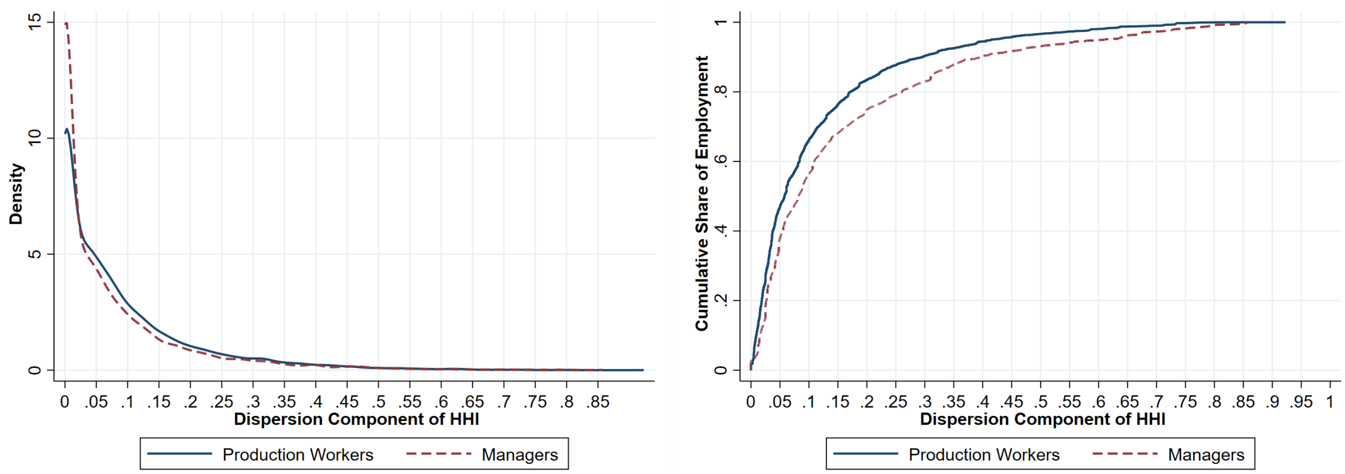
Source: elaboration based on Quadros de Pessoal.

Note: the table reports the (employment weighted) mean, the 10th, 25th, 50th, 75th, and 90th percentiles of the number of establishments across local labor markets by occupation.

for managers. For both occupations, around two-fifths of the observed level of market concentration stems from the number of firms competing in the labor market (see Table 2). Table A.6 reports some distributional moments of the number of establishments across local labor markets. The average production worker works in a market with almost two times as many establishments as the average manager. This fact indicates that only a proportion of establishments find management necessary to organize their production in addition to what is carried out by the entrepreneur. Other things being equal, this observation implies lower market payroll concentration for production workers relative to that of managers.

Fact #3: Greater variance in payroll shares in the average market of managers. The remaining three-fifths of the observed level of concentration stems from the dispersion in payroll shares across establishments relative to the symmetric case (see Table 2). Figure A.8a plots the distribution of the dispersion component of the HHI across local labor markets for each occupation. The distribution of the dispersion component of the HHI is very similar for both production workers and managers. However, looking at the distribution of employment across these market types, we find that fewer production workers work in markets with relatively high levels of payroll dispersion. As a result, the employment-weighted average dispersion level equals 0.15 for managers and 0.11 for production workers. Hence, employment sorting across markets implies that the average manager works in a market with a higher payroll dispersion than the average production worker.

Figure A.8: Distribution of the Variance in Payroll Shares by Occupation



(a) Probability Distribution

(b) Cumulative Share of Employment

Source: Elaboration based on QP.
Note: the graph plots the kernel density of the component of the payroll HHI related to the dispersion in payroll shares to the symmetric market with identical firms. It displays the density for managers (red) and production workers (blue). In addition, the graph plots the CDF of employment with respect to the aforementioned variable.

C Derivations

C.1 Labor Supply

C.2 Labor Demand

C.3 Labor Market Concentration and Markdowns

C.4 Equilibrium

C.5 Quantification of Labor Supply Elasticities

This section explains in detail the quantification of the parameters determining the structural labor supply elasticities: (η_o, θ_o).

C.5.1 Estimation of the Reduced-Form Elasticities

Within-market elasticity

To estimate the within-market *reduced-form labor supply elasticity*, we use establishment-level wages and employment. As both variables are jointly determined in equilibrium, we need to leverage an exogenous change in the labor demand. To this end, we develop a framework with components similar to the one in Ahlfeldt et al. (2022).

We compute total employment $L_{i,j,o,t}$ and average hourly real wages $w_{i,j,o,t}$ in occupation o for each establishment i in location j at period t. It is worth stressing that, unlike Ahlfeldt et al. (2022), we do not need to impute working hours because our database already provides this information. For each occupation o, we separately estimate the following regression:

$$log w_{i,j,o,t} = \beta \; log L_{i,j,o,t} + \mu_{j,o,t} + \upsilon_{i,j,o,t}, \tag{22}$$

where $\beta = 1/\epsilon$ is the inverse of the reduced-form labor supply elasticity. We include location-time fixed effects to isolate any time-varying shock in a given local labor market. Still, the error $\upsilon_{i,j,o,t}$ may capture establishment-time-specific shocks to labor demand and supply, which may be correlated with establishment size. Thus, to consistently estimate β, we use the standard shift-share approach to simulate labor demand shocks that go back to Bartik (1991) and are formalized more recently through the exogeneity of the shares (Goldsmith-Pinkham et al., 2020) or the exogeneity of the shifts (Borusyak et al., 2022). The intuition of the instrument is that we exploit national trends in employment to predict establishment-level labor demand shocks. More concretely, we combine local shares and aggregate shifts to employment as follows:

$$\hat{L}_{i,j,o,t} = \underbrace{\frac{L_{i,j,o,2007}}{\sum_i L_{i,s,o,2007}}}_{\text{Firm's Employment Share in Sector } s} \times \underbrace{\sum_i L_{i,s,o,t}}_{\text{National Employment in Sector } s} . \tag{23}$$

We set 2007 as the initial year for the share component. Then, we measure the shares as the employment in occupation o in a given establishment located in j at t ($L_{i,j,2007,o}$) over the national level of employment for that occupation in sector s in 2007 ($L_{i,j,2007,o}$). Recall that our definition of local labor market j implicitly includes a sector, as it is the intersection between sector (s) and municipality (r) given an occupation o. We multiply this by the total employment of a given sector and occupation ($\sum_i^I L_{iso}$) every year after 2007 to predict current establishment employment according to national trends and initial shares. With the modified employment, we estimate Equation (22) by an instrumental variable with $\hat{L}_{i,j,o,t}$ as an instrument for $L_{i,j,o,t}$. In this estimation, we assume that the instrument is unrelated to unobserved constant or time-varying characteristics that affect specific establishments within the same industry, location, and year.

Table C.1 shows the IV estimates of the reduced-form elasticities by occupations. We find that managers have a smaller labor supply elasticity than production workers, indicating that managers are less responsive to wage changes. Overall, our estimates are in the range of the literature. Most estimates based on inverse methods, i.e., estimating the inverse labor supply elasticity in the baseline specification, find estimates around 5.24 (Sokolova and Sorensen, 2021). Quantifying the model to county-level data in the U.S., Monte et al. (2018) finds a labor supply elasticity of 3.3. Using municipality-level German data, Ahlfeldt et al. (2015) find a labor supply elasticity of 5.5.

Table C.1: Within-Market Firm Substitutability Regression

	Production Workers		Managers	
Dependent Variable: Log Wage	**OLS**	**IV**	**OLS**	**IV**
Log employment	0.0500***	0.0623***	0.1639***	0.1645***
	(0.0002)	(0.0005)	(0.0005)	(0.0017)
Market-Year FE	Yes	Yes	Yes	Yes
Observations	2,580,623	459,960	1,036,216	119,165
Implied Elasticity ($1/\beta$)	19.96	16.05	6.10	6.08
Inferred within-market substitutability (η_o)	19.96	16.05	6.10	6.08

Note: The Table reports the estimates from regressing equation (17) by OLS and IV for each occupation. Standard errors in parentheses. Source: 2002-2016, QP.

Across-market elasticity

To estimate the across-market *reduced-form labor supply elasticity* we use municipality-level wages and employment. The specification, in this case, takes the following form:

$$\text{Log}\, w_{m,o,t} = \beta\, \text{Log}\, L_{m,o,t} + \alpha_{m,o} + e_{m,o,t}, \tag{24}$$

where $w_{m,o,t}$ is the average wage in municipality m, $L_{m,o,t}$ is the total employment in municipality m, and $\alpha_{m,o}$ are municipality fixed effects for each occupation. Because employment and wages are jointly determined in equilibrium, we use the following shift-share instrument for $L_{m,o,t}$:

$$\hat{L}_{m,o,t} = \sum_s \left(\underbrace{\frac{L_{i,m,s,o,2007}}{\sum_i L_{i,m,s,o,2007}}}_{\text{Industry-Municipality Share}} \times \underbrace{\sum_i L_{i,s,o,t}}_{\text{National Employment in Sector } s} \right). \tag{25}$$

The intuition for the instrument is similar to before, but now we exploit across-municipality variation, instead of establishment-level variation, over time that stems from national employment shocks to sectors. To explain why the instrument might be valid, we argue that multiple shifts to employment by sector, at the national level, are unrelated to local economic conditions. Hence, in this setup, the national employment trends by sector adjust the local labor demand exogenously.

In terms of results, Table C.2 shows the estimates from regression 24. We find that the elasticity is lower for managers (1.4) than for production workers (2.53). Suggesting that, according to the definition of our model, it is more costly for managers to move across markets than for production workers.

Table C.2: Across-Market Firm Substitutability Regression

	(1)	(2)
	Production Workers	Managers
Log employment	0.433***	1.008***
	(0.027)	(0.078)
Municipality FE	Yes	Yes
Observations	1,946	1,946
Implied Elasticity ($1/\beta$)	2.31	0.99
Inferred across-market substitutability (θ_o)	2.58	0.70

Note: The Table reports the estimates from equation (18) with IV. Confidence intervals at the 95% level. Standard errors are clustered at the municipality level. Source: 2007-2016, QP.

C.6 Estimation of the Mass Layoff Shocks

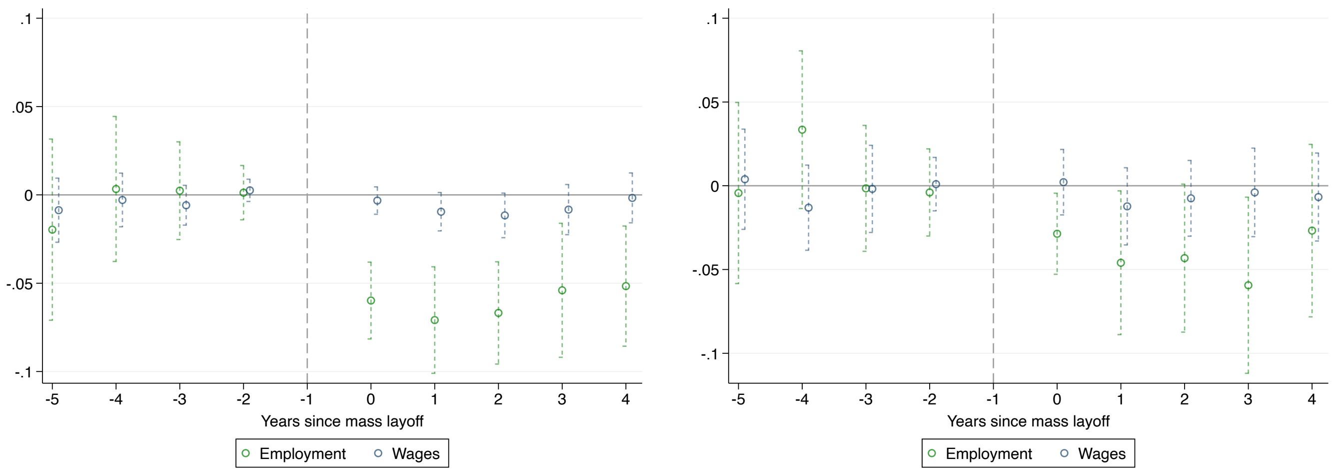
The mass layoff definition we use is a drop of 100 workers in any establishment for two consecutive years between 2004 to 2016. Moreover, we also require the establishment to account for at least one percent of municipality's employment. We make this last requirement to consider sizable shocks within the local economy. To quantify the impact of mass-layoff shocks at the municipal level, we use an event-study specification that takes the following form:

$$y_{mt} = \xi + \sum_{k=-5}^{5} \beta_k \mathbf{1}\{k = t - g\} + \gamma_m + \gamma_t + X'_{it}\theta + \epsilon_{mt}. \tag{26}$$

Here, the year that a mass layoff occurs in an establishment of a municipality is g, the years are t, and the event time indicators are k.[12] Year fixed effects (γ_t) and municipality fixed effects (γ_m) control for unobserved constant characteristics across all municipalities and within municipalities, respectively. The covariates X_{it} control for baseline characteristics regarding the size and structure of the cities interacted with time to flexibly control for time-varying variables. More precisely, these variables are the log of municipality's employment, the share of manufacturing employment, the share of highly educated workers, the share of male workers, and the share of young workers. The parameters of interest are β_k, which come from k event time dummy variables. These dynamic treatment effects measure the effect on y relative to an omitted period, which is when $k = -1$. We use Callaway and Sant'Anna (2021) estimator to control further for heterogenous treatment effects across cohorts and aggregate all results across cohorts for the main coefficient shown in the main text. Figure C.1a shows the dynamic pre- and post-treatment coefficient on employment and wages following a mass layoff shock. Importantly, pre-treatment coefficients are around zero, indicating that treated and control groups were on a similar path before treated municipalities were hit by a mass layoff. Overall, the most negative significant results happen on the employment margin, not on the wage margin, for both types of workers.

[12] We define the cohorts of treated municipalities from the first time, in our sample period, they had a mass layoff. Other mass layoffs may happen in the same municipality later on.

Figure C.1: Event Study Estimates by Occupation



(a) Production workers

(b) Managers

Note: These coefficients plot the estimated event study estimates using Callaway and Sant'Anna (2021) estimator. We exploit 72 events of mass layoffs across time in our sample period. We identify mass layoffs as a drop of 100 workers in a given establishment for two consecutive years. Source: QP, 2004-2016.

C.7 Quantification of Firm Distribution