Monopsony Power and Firm Organization*

Álvaro Jáñez[a], Lukas Delgado-Prieto[b]

[a]Stockholm School of Economics
[b]University of Oslo

August 20, 2025
[Click here for the latest version]

Abstract

Labor market competition drastically differs for production workers and managers. We extend a general equilibrium model of oligopsony with minimum wages to include firm organization and production complementarities across occupations. Using administrative data from Portugal, we estimate the model and validate it against quasi-experimental evidence on oligopsony and minimum wage effects. Relative to the efficient economy, monopsony reduces employment and wages, particularly for managers at the most productive firms. Managers' monopsony alone accounts for one-fifth of employment and wage losses for production workers. Welfare losses from monopsony are 3.4 and 2.4 percent for managers and production workers, respectively. Production workers bear smaller losses because they often work in markets with more competitor firms, view firms as closer substitutes, and are more affected by the minimum wage.

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 Juan J. Dolado, Andrés Erosa, Luisa Fuster, Carlo Galli, Matthias Kredler, and Emircan Yurdagul for many insightful discussions. We are thankful for comments from Manuel Amador, Kyle Herkenhoff, Timothy J. Kehoe, and Simon Mongey, 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 also thank João Pereira dos Santos for his invaluable support to access the data. Álvaro Jáñez gratefully acknowledges support from the Ministerio de Ciencia e Innovación through research grants PRE2019-088620, the AEI (Spain) through research grant PID2019-110779GB-I00, and Universidad Carlos III through the scholarship for Master's study. Lukas Delgado-Prieto acknowledges the financial support from the Ministry of Science and Innovation in Spain through research grant PDI2019-108144GB-I00. Corresponding author: Stockholm School of Economics, Sveavägen 65, 113 83 Stockholm, Sweden. E-mail: alvaro.janez.garcia@hhs.se.

1 Introduction

Imperfect labor market competition enables firms to offer lower wages, leading to lower employment, labor misallocation, and welfare losses. The number of competing firms, the ease of substituting between employers, and institutions such as minimum wages shape firms' wage-setting power. These factors vary across occupations and, when these are complementary inputs, amplify the negative effects from monopsony. For instance, in Portugal, half as many firms compete for managers as for production workers in the typical local labor market, as only some firms delegate decision-making to managers.[1] Since limited competition incentivizes firms to hire fewer managers, firms also lose production worker teams under their supervision. This paper shows that the heterogeneity and interactions in monopsony between these two occupations matter when studying the welfare effects of monopsony power.

To that end, we develop a general equilibrium model with firm-occupation-specific monopsony power and estimate it using administrative data from Portugal. We find that, in the most equal scenario, removing monopsony leads to a consumption equivalent welfare gain of 3.4 percent for managers and 2.4 percent for production workers relative to the benchmark. Removing managers' monopsony alone accounts for one-fifth of the earnings gains of production workers through production complementarities, but reduces production workers' welfare by rebating profit losses. In terms of policy, even a social welfare-maximizing occupation-specific minimum wage captures less than one-tenth of the gains from efficiency.

To arrive at these conclusions, we extend a general equilibrium model of oligopsony to incorporate minimum wages and firm organization. The economy features a representative household for each occupation and a continuum of local labor markets, each with a finite number of firms. Households choose consumption, capital, and labor supply to each individual firm for their respective occupations, viewing firms within the same and across distinct markets as imperfect substitutes with occupation-specific degrees of substitutability. Firms exogenously differ in terms of productivity and the local labor market they inhabit. Each firm's organizational decision involves adopting a single- or two-layer organizational structure. Single-layer firms hire only production workers, while two-layer firms add a management layer and decide

[1] We exclude CEOs and most managers are supervisors, team leaders, or middle managers. For each occupation, we define a local labor market as the combination of a municipality and a two-digit industry.

how many managers and production workers to hire. Managers enable firms to expand their workforce, yet it comes at higher costs, and thus, only the most productive firms adopt a two-layer structure. Regarding monopsony power, firm-occupation-specific markdowns arise from differences in (i) sorting across firms of distinct size, (ii) firm substitutability, and (iii) exposure to the statutory minimum wage.

We estimate the model using the Simulated Method of Moments (SMM) to fit moments of firm organization, wages, and market concentration in Portugal. The firm substitutability parameters are key for the amount of monopsony power in each occupation. We exploit the correlation between employment and wages at the establishment level, controlling for market unobserved heterogeneity, to calibrate the within-market firm substitutability parameters. Moreover, we adopt an indirect inference approach to estimate the across-market firm substitutability parameters from plausibly exogenous labor demand changes at the municipality level, which we generate with a Bartik-type instrument that exploits the exposure of each municipality to national sector value-added trends since the Great Recession.

We validate the model by comparing the model predictions to cross-sectional moments and quasi-experimental evidence about the presence of oligopsony, firm organization, and the effects of minimum wage policies. The model reproduces the empirical distribution of employment, wages, and the span of control of managers across firms. It also matches average moments of market concentration across occupations. Moreover, the model quantitatively replicates two reduced-form experiments. First, it matches quasi-experimental evidence on the pass-through of idiosyncratic demand shocks to wages in Portugal (Garin and Silvério, 2024; henceforth GS). The model-generated pass-through estimate of 0.17 is not significantly different from GS's estimate, and the model also reproduces higher pass-through in more concentrated markets. Second, the model aligns with employment effects of minimum wage policies from a comprehensive set of developed countries (Dube and Zipperer, 2024). Specifically, the model produces most of the own-wage-elasticity (OWE) estimates, which measure the employment and mean wage responses to minimum wage changes, within the range of empirical estimates, and correctly captures that, on average, raising the statutory minimum wage generates fairly modest employment losses and increases earnings.

To quantify the effect of monopsony power on aggregate outcomes, we compare the benchmark equilibrium with a counterfactual efficient economy where we exogenously set wages to the marginal product of labor. In the benchmark economy, we estimate that the average manager and production worker bear a wage markdown of 31.9 and 16.0 percent, respectively. Firms generally exert wider wage markdowns on managers because managers (i) often work in markets with fewer competitor firms, (ii) the minimum wage is more likely to be binding for production workers, and (iii) have lower across and within market elasticities. Regarding the lower elasticities of managers, our result is consistent with evidence on the lower labor supply elasticity of high-wage workers (Diamond, 2016; Langella and Manning, 2021; Bachmann et al., 2022; Goolsbee and Syverson, 2023).

We derive four main results when comparing the efficient to the benchmark economy. First, employment and wages increase, along with the concentration of employment at the most productive firms, especially for managers. Second, part of this increase in employment concentration arises from rising wages making management delegation unprofitable for medium-productivity firms, decreasing the share of two-layer firms by 10.9 percent in the efficient economy. This organizational change increases manager concentration at the most productive firms, which also expand their hiring of production workers and production. Third, removing managers' monopsony alone contributes significantly to efficiency gains due to production complementarities. Specifically, it accounts for nearly one-fifth of the increase in employment, wages, and employment concentration of production workers, while explaining about three-fifths of the output gains in the efficient relative to the benchmark economy. Fourth, the efficient economy provides a social welfare gain of 2.7 percent in consumption equivalent units relative to the benchmark due to higher earnings and despite worker reallocation and profit losses. The welfare effects of monopsony power on each occupation depend on their respective profit share. Considering a range of profit shares between equal shares and population shares, we find that while managers always enjoy welfare gains, production workers' welfare only rises when they bear sufficiently low profit losses. Under equal profit shares, the welfare gain is 3.4 percent for managers and 2.4 for production workers.

To assess the implications for the design of optimal minimum wage policies to alleviate monopsony power, we analyze whether an occupation-specific, rather than a single statutory minimum wage, is more effective at capturing the gains from the efficient economy.[2] We find that the optimal single statutory minimum wage, which maximizes social welfare with population weights, captures less than 10 percent of the welfare gains from an efficient economy. Despite the presence of occupation-specific monopsony power, an optimal occupation-specific minimum wage only slightly improves upon the optimal single minimum wage, as these minimum wages inevitably bind first for low-productivity firms within each occupation, where monopsony power is relatively weak. Moreover, the optimal manager-specific minimum wage is relatively lower than that for production workers, despite stronger monopsony power over managers. This occurs because social welfare depends primarily on production worker outcomes, so a low manager-specific minimum wage mitigates managerial job losses and the associated decline in labor demand and welfare for production workers.

Literature. This paper contributes to the literature on oligopsonistic labor markets and how this affects the overall economy (Bhaskar et al., 2002; MacKenzie, 2021; Berger et al., 2022; Jarosch et al., 2023; Azkarate-Askasua and Zerecero, 2024; Deb et al., 2024). Using models where monopsony power arises from firm granularity and imperfect firm substitutability, they study the effect of monopsony on wages, efficiency, and welfare. Our main theoretical contribution to this literature is to study the effect of monopsony power on these outcomes through the organization of work within firms. The distinctive mechanisms in our model are that (i) firms make organizational decisions that endogenously contribute to markdown heterogeneity, and (ii) worker types are complementary in production. Quantitatively, we show that these contributions are key to understanding how monopsony power affects workers' outcomes, compresses the firm size distribution, and reduces welfare.

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). We build our framework on Berger et al. (2025), which studies the effect of minimum wages on efficiency and welfare in an oligopsonistic environment with firm and worker heterogeneity. Our main contribution is to allow for occupation-specific markdowns and imperfect substitutability across worker types in production. This matters for adding two findings to this literature. First, we rationalize

[2] Occupation-based minimum wages are implemented in Australia, with its Modern Awards legislation, and are common in many European countries, whose collective contracts set occupation-specific wage floors.

that minimum wages affect the employment and wage distribution of managers by affecting production workers. Second, we show that an optimal occupation-based minimum wages does not significantly improve upon an optimal single statutory minimum wage.

This paper also contributes to the literature on production organization (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). Contemporaneously to our work, Lawson et al. (2023) studies the impact of minimum wages on productivity through firm organization in a perfectly competitive framework. To the best of our knowledge, we are the first to incorporate monopsony power in a general equilibrium model with managerial delegation choices. This adds two contributions to this literature. First, delegation choices help to explain the degree of monopsony power over managers and production workers. Second, we show that lower labor market competition depresses managers' wages, especially in large firms, and therefore increases the share of firms that delegate tasks to managers.

We also contribute 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 monopsony power compresses the firm size distribution mainly by making medium-productivity firms inefficiently large, as these firms have an inefficiently high share of managers.

Finally, we relate to the literature quantifying labor supply elasticities across worker types, which shows that elasticities decline with educational attainment (Diamond, 2016), are lowest among top earners (Langella and Manning, 2021), are lower for workers in nonroutine cognitive than in routine or manual tasks (Bachmann et al., 2022), and fall with tenure in higher education labor markets (Goolsbee and Syverson, 2023). Consistent with these results, we find that managers exhibit lower labor supply elasticities than production workers.

2 Model

This section presents a general equilibrium model that incorporates firm organization, oligopsonistic labor markets, and minimum wages. The model considers two permanent occupations, managers and production workers, each with heterogeneous labor disutility costs and substitutability across firms. For each occupation, there is a household that chooses consumption, the capital stock, and the labor supply to each firm. Firms are heterogeneous in productivity and the local labor market they inhabit. Regarding their organization, firms have a layer of production workers and choose whether to add a management layer. Moreover, they choose how much capital to rent and the number of workers in each layer. Firms have monopsony power and face a minimum wage when making employment choices, where wage markdowns are firm-occupation-specific.

2.1 Environment

Agents. The economy consists of two households, indexed by their permanent occupation type $o \in \{w, m\}$, and a continuum of firms. Households are ex-ante heterogeneous in their disutility of labor, which depends on the aggregate labor supply and the allocation of labor across firms. Firms are heterogeneous in two dimensions. First, they inhabit distinct locations $j \in [0, 1]$, each with a finite number of firms indexed by $i \in \{1, \dots, M_j\}$. Second, they differ in productivity z_{ijt}, drawn from a standard log-normal distribution with standard deviation σ_z.

Goods and Technology. Firms combine capital and labor to produce a tradable good in a perfectly competitive national market whose price we normalize to one. We assume there are two types of labor: production workers (n_w) and managers (n_m). Production workers are essential for production, while managers are optional. Specifically, each firm chooses between two types of organizations, which vary in the number of layers, $\ell \in \{1, 2\}$. The production technology of single-layer firms ($\ell = 1$) is given by:

$$y(z, 1) \;=\; z^{w}_{ijt}\Big(k^{1-\gamma}_{ijt}\, n^{\gamma}_{ijwt}\Big)^{\alpha}, \quad \gamma \in (0, 1),\; \alpha > 0. \tag{1}$$

The total factor productivity (TFP) of single-layer firms is the product of the idiosyncratic component and a layer-specific shifter, $z^{w}_{ijt} = \bar{z}_w z_{ijt}$. This technology captures the essential features of the standard Lucas (1978) model, incorporating productivity heterogeneity and allowing for diminishing returns to scale. The degree of returns to scale α governs how much firms can expand production by increasing the number of production workers. Alternatively, firms may add a managerial layer ($\ell = 2$) and produce according to:

$$y(z, 2) = z^m_{ijt}\, n^{(1-\alpha)\alpha}_{ijmt} \Big(k^{1-\gamma}_{ijt}\, n^{\gamma}_{ijwt}\Big)^{\alpha}. \tag{2}$$

Managerial delegation enhances the productivity of capital and production workers through two distinct channels. First, managerial labor increases the marginal productivity of other inputs at any given TFP level. Second, we allow for layer-specific TFP differences, $z^m_{ijt} = \bar{z}_m z_{ijt}$. Overall, this technological specification enables tractable model quantification while capturing the main trade-off from the firm organization literature: adding a managerial layer allows firms to manage larger workforces but imposes higher equilibrium costs (e.g., see Garicano, 2000; Garicano and Rossi-Hansberg, 2006). Beyond determining returns to scale, the parameter α governs the managers' span of control. This approach naturally embeds both concepts within a single parameter that reflects the technological constraints of expanding production through additional production workers. We will show shortly that the quantification based on returns to scale aligns with evidence on the span of control.

Households. Each household type $o \in \{w, m\}$ chooses the measure of workers to supply to each firm n_{ijot}, the capital stock in the next period K_{ot+1} and consumption of each good c_{ijot} to maximize their utility:

$$\mathcal{U}_{ot} = \max_{\{n_{ijot}, c_{ijot}, K_{ot+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \left[\mathbf{C}_{ot} \;-\; \varphi_o \frac{\mathbf{N}_{ot}^{1+\frac{1}{\phi}}}{1+\frac{1}{\phi}}\right], \tag{3}$$

subject to the household's budget constraint:

$$\mathbf{C}_{ot} + [K_{ot+1} - (1-\delta)K_{ot}] \;=\; \int_0^1 \sum_{i=1}^{M_j} w_{ijo} n_{ijo}\, dj \;+ R_t K_{ot} + \kappa_o \Pi_t, \tag{4}$$

where we define the aggregate consumption and labor supply indexes as:

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

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

The parameter ϕ stands for the aggregate Frisch elasticity of households, φ_o is an occupation-specific labor disutility shifter, and B_{jo} is an occupation-specific location amenity shifter, and κ_o stands for the occupation-specific fraction of profits rebated to the household. We assume that consumption goods are perfectly substitutable, but households view firms as imperfect substitutes in terms of non-wage characteristics (see, e.g., Berger et al., 2022). In addition, we extend the model to allow individual firms to face *occupation-specific* upward-sloping labor supply curves with two elasticities of substitution θ_o and η_o. The parameter θ_o regulates the degree of substitutability of firms in distinct markets and, thus, captures the costs of moving across markets or idiosyncratic tastes for the market. If these costs decrease ($\theta_o \uparrow$), workers find it easier to substitute firms across markets and become more responsive to market wage differentials. The parameter η_o regulates the degree of substitutability of firms within the same market, thus capturing features such as commuting costs, search costs, or idiosyncratic tastes for the firm. As 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. When $\eta_0 > \theta_0 > 0$, the household $o \in \{w, m\}$ perceives firms within the same market as closer substitutes than firms across different markets. Consequently, larger firms hinder labor reallocation to other firms by reducing the number of alternative employers within the same market, forcing workers to seek less substitutable firms in other markets. This limited substitutability reduces workers' responsiveness to individual firm wage policies, thereby conferring greater monopsony power to larger firms.

For each occupation $o \in \{w, m\}$, the first order necessary conditions of the utility maximization problem imply that the supply of capital is infinitely elastic and the aggregate labor supply is given by:

$$R_t = \frac{1}{\beta} - (1 - \delta), \tag{5}$$

the aggregate supply of labor is given by:

$$\mathbf{N}_{ot} = \left(\frac{\mathbf{W}_{ot}}{\varphi_o} \right)^{\phi}, \tag{6}$$

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

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

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

$$\mathbf{w}_{jot} := \left[\sum_{i\in j} w_{ijot}^{1+\eta_o}\right]^{\frac{1}{1+\eta_o}} \qquad \mathbf{W}_{ot} := \left[\int_0^1 \left(B_{jo}\mathbf{w}_{jot}\right)^{1+\theta_o} dj\right]^{\frac{1}{1+\theta_o}}. \quad (8)$$

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

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

When deciding on the optimal organizational structure, firms compare their maximum profits as a single-layer and a two-layer organization. Single-layer organizations choose how much capital to rent, k_{ijt}, and the number of production workers to hire, n_{ijwt}. Two-layer organizations additionally choose the number of managers to hire, n_{ijmt}. When making these decisions, both organization take as given the labor supply curves, the labor demand of their competitors within the same market (n^*_{-ijot}), the statutory minimum wage ($\underline{w}$), and the aggregate variables ($\mathbf{W}_{ot}, \mathbf{N}_{ot}$). Formally, single-layer organizations maximize profits:

$$\pi(z,1) \;=\; \max_{n_{ijwt},k_{ijt}} \quad y(z,1) \;-\; R_t k_{ijt} \;-\; w_{ijwt}\Big(n_{ijwt}, n^*_{-ijwt}, \mathbf{N}_{wt}, \mathbf{W}_{wt}\Big)n_{ijwt}, \quad (10)$$

subject to:

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

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

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

In addition, the profit maximization problem of two-layer organizations is given by:

$$\pi(z,2) \;=\; \max_{n_{ijwt},n_{ijmt},k_{ijt}} \quad y(z,2) \;-\; R_t k_{ijt} \;-\; \sum_{o\in\{w,m\}} w_{ijot}\Big(n_{ijot}, n^*_{-ijot}, \mathbf{N}_{ot}, \mathbf{W}_{ot}\Big) n_{ijot}, \tag{11}$$

subject to:

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

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

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

The first order necessary condition for capital implies that the interest rate equal the marginal product of capital:

$$\frac{\partial y(z,\ell)}{\partial k} \;=\; R_t, \quad \forall \ell \in \{1,2\}.$$

For each each occupation $o \in \{w,m\}$, the presence of the statutory minimum wage implies that the solution for the labor demand has three cases. First, the minimum wage is not binding. Second, the minimum wage is binding, and labor demand equals the labor supply curve. Third, the minimum wage is binding and labor supply exceeds labor demand. We summarize the system of first-order conditions for each case as follows.

Case I: The minimum wage is not binding. The marginal cost of labor equals its marginal product at the optimal employment level:

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

When the structural elasticity is positive, $\varepsilon_{ijot} > 0$, Equation (12) implies that workers earn wages below their marginal productivity (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. Hence, there is a wedge between wages and the marginal product of labor $\mu_{ijot} < 1$. In addition, the oligopsonistic market structure implies that the structural labor supply elasticity depends on the payroll share of the firm:

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

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

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

The model explicitly tells apart the potential forces shaping wage dispersion across occupations. The first source of dispersion comes from differences in marginal productivity, which partly depends on organizational choices. The second source of dispersion comes from firm-occupation-specific markdowns due to different (i) sorting across firms of distinct size (s_{ijot}), (ii) firm substitutability (η_o, θ_o), and (iii) exposure to the statutory minimum wage, which we describe in the next cases.

Case II: The minimum wage is binding, and labor supply equals labor demand. The minimum wage is binding and is below the efficient wage level where the labor supply curve intersects the marginal product curve. In this case, firms pay the minimum wage, and the markdown is the ratio between the minimum wage and the marginal product, with the level of employment given by the labor supply curve evaluated at the minimum wage:

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

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

Case III: The minimum wage is binding, and labor supply excess labor demand. The minimum wage is binding and is above the efficient wage level where the labor supply curve intersects the marginal product curve. In this case, firms pay a wage that is equal to both the minimum wage and marginal product, with the employment level given by the marginal product, and firms potentially face an excess of labor supply:

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

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

Equilibrium. Given the statutory minimum wage $\underline{w}$, the steady state general equilibrium of the model is a set of organizational structures $\{\ell^*_{ij}\}_{\forall ijt}$, aggregate disutilities of labor supply $(\mathbf{N}^*_{wt}, \mathbf{N}^*_{mt})$, and employment levels $\{n^*_{ijwt}, n^*_{ijmt}\}_{\forall ij}$ such that:

1. *Households*: households choose aggregate supply of labor, $\mathbf{N}^*_{ot}$, labor supply to each individual firm $\{n^*_{ijot}\}$, and capital, K_{ot}, to maximize utility. That is, Equations (5)-(7) hold $\forall o \in \{w, m\}$, $\forall j \in [0,1]$, and $\forall i = \{1, \ldots, M_j\}$.

2. *Firms*: firms optimally choose the organizational structure, ℓ^*_{ijt}, capital, k^*_{ijt}, and the number of workers to hire in each occupation, n^*_{ijot}. That is, Equations (11)-(16) hold $\forall o \in \{w, m\}$, $\forall j \in [0,1]$, and $\forall i = \{1, \ldots, M_j\}$.

3. *Market Clearing*: all markets clear,
   - Output: $\int_o^1 \sum_i^{m_j} y^*_{ijt}\, dj \;=\; \sum_{o\in\{w,m\}}(C^*_{ot} + \delta K^*_{ot})$.
   - Capital: $\left.\frac{\partial y_{ijt}}{\partial k}\right|_{k^*_{ijt}} \;=\; \frac{1}{\beta} - (1-\delta)\;\; \forall j \in [0,1]$ and $\forall i = \{1, \ldots, M_j\}$.
   - Labor: labor supply and demand are given by Equations (12) and (15) for firms in Cases I and II. For firms in Case III, households supply the labor demand n^*_{ijot} given by Equation (16).

Note that the equilibrium considers market clearing in the presence of minimum wages. To handle non-market-clearing wages, we solve the equilibrium using a shadow wages approach as in Berger et al. (2025). This approach considers that households perceive a lower wage than the minimum wage for firms in Case III, which implies that the excess labor supply at the minimum wage is reallocated towards other firms (see Appendix D). Appendix F discusses the interaction between firm organization and monopsony power in equilibrium.

3 Quantification of the Model

The quantification of the model parameters proceeds in three steps: (i) we exogenously calibrate the minimum wage and Frisch elasticity; (ii) we endogenously calibrate the discount factor, several technological parameters, and the within-market firm substitutability parameters; and (iii) we jointly estimate the remaining model parameters, including the across-market firm substitutability parameters, using the Simulated Method of Moments (SMM) approach. The firm substitutability parameters are key for the amount of monopsony power. We set the within-market substitutability parameters to match the slope between wages and employment controlling for market-year fixed effects, and we use an indirect inference approach to estimate the across-market substitutability parameters from plausibly exogenous changes in municipality's labor demand.

3.1 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, hourly wages, occupation, industry, and geography for all private firms based in Portugal with at least one worker. We use anonymised firm identifiers to link these data to balance sheet data from the *Sistema de Contas Integradas das Empresas* (SCIE). Our sample period covers from 2010 to 2016. We explain here the main aspects of the sample and relegate the details to Appendix B.

We assign workers to each occupation following a hierarchical classification similar to Caliendo et al. (2020). By Portuguese law, firms must assign workers to hierarchic categories that allow us to distinguish between two layers within each firm (see Tables A.1 and A.2). We exclude CEOs and assign middle managers, supervisors, team leaders, and top managers to the management layer. The distinctive feature of managers is that they guide groups of employees in their tasks. We group the remaining categories as production workers, which range from non-skilled to higher-skilled professionals.[3]

[3] According to occupation transitions, these broad categories represent a persistent occupational state. Figure A.5 shows that most workers remain within the same category after changing to another establishment.

Next, we define a labor market for each occupation, jo, based on their geography (municipality) and industry (2-digit NACE), capturing that workers are more attached to their current labor market due to imperfect skill substitutability and costly geographical mobility (Neal, 1995; Kambourov and Manovskii, 2009; Sullivan, 2010; Kennan and Walker, 2011).

Lastly, we use the balance sheet data to compute the capital and labor share. Capital consists of investment properties, tangible fixed assets, and intangible fixed assets. We use an average return on capital of 15 percent (Barkai, 2020) and calculate the capital share as the ratio between the return of the capital stock and value added. Labor income corresponds to all personnel expenses, including employee remunerations and social charges (e.g., pensions or severance payments). The labor share is the ratio between labor income and value added.

3.2 Calibrated parameters

Minimum wage, technology and preferences. Table 1 summarizes the model parameters. We calibrate outside the model the statutory minimum wage and the aggregate Frisch elasticity. We adjust all nominal variables using 2012 CPI and use the 2016 statutory minimum wage, which yields $\underline{w} = 525$. 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. On an annual basis, we endogenously calibrate the discount factor ($\beta = 0.96$) and the depreciation rate ($\delta = 0$) to match a discount and interest rate of 4 percent. We calibrate the parameter governing decreasing returns to scale (α) and the exponent on labor (γ) using the labor and capital share, respectively.

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

$$\log(w^*_{ijo}) \ = \ \frac{1}{\eta_o}\log(n^*_{ijo}) \ + \ \underbrace{\underbrace{\left(\frac{1}{\theta_o} - \frac{1}{\eta_o}\right)\log\Big(\mathbf{n}_{jo}(n^*_{ijo}, n^*_{-ijo})\Big)}_{\text{Effect of payroll share on wages due to } \mathbf{n}_{jo}} + \frac{1}{\theta_o}\log(\mathbf{N}_o) \ + \ \log(\mathbf{W}_o)}_{\text{Common across firms in local labor market } jo}.$$

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 strategic interactions on the labor supply elasticity shuts down when we control for market fixed effects. We use this insight to obtain a theory-consistent estimate of the within-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 control for common labor demand and supply considerations within 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) identifies the within-market elasticity as $\hat{\eta}_o = 1/\hat{\beta}_o$. 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. Increasing the productivity of a firm has two equilibrium effects. First, the labor demand curve shifts up because the marginal productivity rises. Second, the labor supply curve shifts down 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.

To address potentially endogenous firm-level supply shocks in the data, we estimate Equation (17) by IV in our baseline specification. We use a value-added shift-share instrument that predict firms' employment from national sector employment trends and initial firm shares (Severen, 2021; Ahlfeldt et al., 2022):

$$\hat{n}_{ijo,t} = \underbrace{\frac{y_{is(j),2004}}{\sum_i y_{is(j),2004}}}_{\substack{\text{Industry-firm} \\ \text{share}}} \times \underbrace{\sum_i y_{is(j),t}}_{\substack{\text{National value added} \\ \text{in industry } s}} .$$

We restrict the sample to firms paying wages at least five percent higher than the minimum wage of the reference year. This regression implies a within-market labor supply elasticity of 7.8 for production workers and 2.3 for managers (see Table E.3). Thus, we find that production workers are three times as responsive to changes in labor demand as managers.

3.3 Estimated parameters

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 the vector of model moments and its data counterpart. Next, we describe each parameter and its most informative moment in detail.[4]

Labor disutility shifter (ϕ_w, ϕ_m). The most associated moment with the labor disutility shifter of production workers ϕ_w is the average firm size. In the data, the average firm hires 5.3 production workers. For the labor disutility shifter of managers ϕ_m, we include as the most informative moment that about 18.6 percent of all employees are managers.

Efficiency of labor ($\bar{z}_w, \bar{z}_m$). The efficiency of each organization type is informative of wages. Thus, we include as targets the mean monthly wage of 867€ for production workers and the wage gap in mean wages between managers and production workers, which is equal to 0.62 log points.[5]

Dispersion in firm productivity (σ_z). Labor markets are more concentrated in the presence of higher productivity differentials across firms. Thus, we choose the employment-weighted average HHI in production workers' local labor markets, which is equal to 0.20, as the moment most associated with the standard deviation of firm productivity.

Market Amenities (B_{ijw}). We note that only 12 percent of production workers belong to markets with fewer than ten firms, despite these markets representing nearly 65 percent of the total. Thus, we set a common market amenity for production workers in these markets and include this moment as the most informative of market amenities in the estimation. The rationale for using amenities rather than productivity differences is that we would require excessively low productivity in small markets, which would overestimate the share of minimum wage earners in such markets.

Firm Distribution (G). The distribution of the number of firms across markets, $M_j \sim G(\cdot)$, combines a discrete mass at $m_j = 1$ with a Pareto distribution. To estimate these

[4] Appendix E.1 shows the fit of the targeted moments.

[5] We are particularly interested in matching wage differential at the bottom of the distribution. Thus, we restrict the sample to managerial wages below the 90th percentile for this moment.

Table 1: Parameterization

Parameter Value	**Description**	**Value**	**Moment**
Panel I: Exogenous calibration			
ϕ	Aggregate Frisch elasticity	0.50	Berger et al. (2022)
$\underline{w}$	Minimum wage	525	Real minimum wage in 2016
Panel II: Endogenous calibration			
β	Discount factor	0.96	Annual discount rate of 4%
δ	Depreciation of capital	0	Annual interest rate of 4%
α	Decreasing returns to scale	0.55	Labor share of 62%
γ	Exponent on labor	0.82	Capital share of 31%
(η_w, η_m)	Within-market substitutability	(7.82 , 2.32)	Within-market labor supply elasticity
Panel III: SMM Estimation			
A: Preferences			
φ_w	Labor disutility: workers	122	Average firm size
φ_m	Labor disutility: managers	1.4	Share managers
B: Firm Organization			
$\bar{z}_w$	Worker efficiency	1,062	Mean wage of prod. workers
$\bar{z}_m/\bar{z}_w$	Managerial efficiency	2.1	Wage gap managers vs prod. workers
σ_z	Std. Dev. firm TFP	0.7	Weighted mean HHI prod. workers
C: Market Characteristics			
B_{ijw}	Amenities in small markets	0.7	Share workers in markets $M_j \leq 10$
$G(\cdot)$	Firm distribution		Mean, variance, and mass single-firm
D: Firm Substitutability			
(θ_w, θ_m)	Across-market substitutability	(2.4 , 1.0)	Across-municipality labor supply elasticity

Source: The Table reports the quantification of model parameters. Panel I reports the parameters that we calibrate outside the model. Panel II reports the estimated parameters using the SMM approach.

parameters, the most associated moments include that 29 percent of markets have just one firm, the average market has 17.4 firms, and the standard deviation in the number of firms equals 59.9.

Across-market substitutability (θ_w, θ_m). The across-market firm substitutability parameters govern how greater market productivity translates into more employment. When firm substitutability is high, 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 equation:

$$\text{Log}\, w_{mo,t} = \gamma \,\text{Log}\, n_{mo,t} + \alpha_{mo} + e_{mo,t}, \tag{18}$$

where $w_{m,o,t}$ is the mean wage in municipality m for occupation o in period t, $n_{m,o,t}$ is total employment in that municipality, and $\alpha_{m,o}$ are municipality fixed effects. The main threat to identification is that employment and wages in a municipality may vary over time due to changes in labor supply. To overcome this problem, we use a value-added shift-share instrument for employment (Lamadon et al., 2022):[6]

$$\hat{n}_{mo,t} = \sum_s \left(\underbrace{\frac{y_{imso,2004}}{\sum_i y_{imso,2004}}}_{\substack{\text{Industry-municipality} \\ \text{initial share}}} \times \underbrace{\sum_i y_{iso,t}}_{\substack{\text{National value-added} \\ \text{in sector } s}} \right). \tag{19}$$

This instrument predicts employment in each municipality as the sector-wide value-added growth interacted with the past concentration of that sector's value added across municipalities. 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 initial exposure to them, which plausibly represent exogenous changes in the municipality's labor demand. We also restrict the sample to municipalities with a mean wage higher than 5 percent of the minimum wage in the reference year. We impose this restriction to exclude municipalities where a high share of firms pay wages close to the

[6] 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.

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 E.2 reports the IV results for each occupation. The implied coefficients are 2.5 for production workers and 1.0 for managers.

To replicate this regression in the model, we randomly assign markets to each municipality to 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 $\sigma = 0.05$. Moreover, we restrict the sample of municipalities to those with a mean wage higher than 5 percent of the model minimum wage in both periods. 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 2.5 for production workers and 1.0 for managers.

Given the importance of both the within- and across-market firm-substitutability parameters in the distribution of wage markdowns, Appendix E.3 shows that our findings are consistent with and fall within the range of labor supply elasticity estimates in the literature.

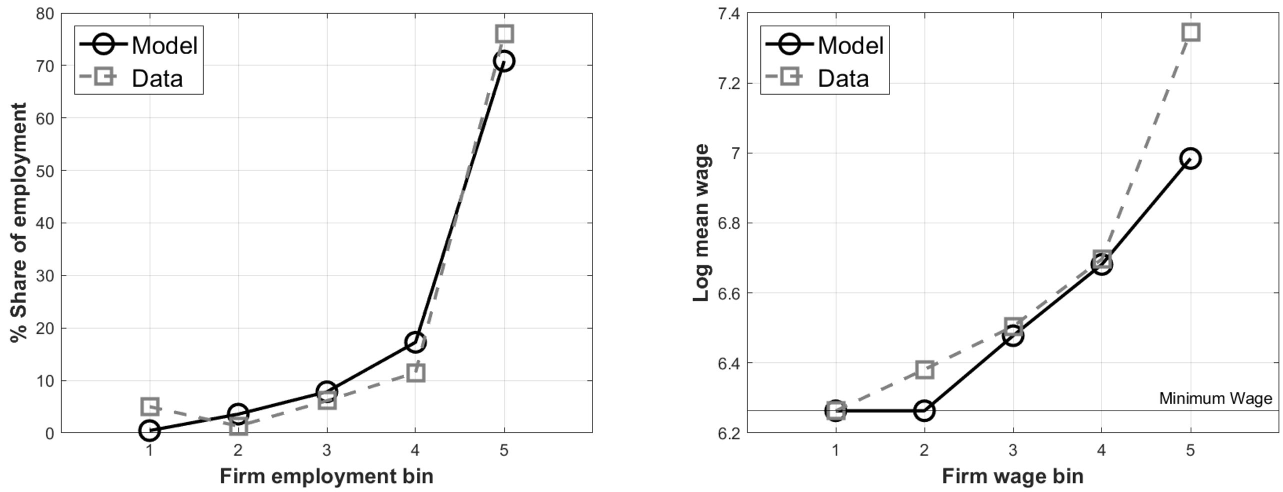
4 Validation of the model

We compare the model predictions to empirics that are informative about the presence of oligopsony, firm organization, and the effects of minimum wages. The model closely matches moments of wages, employment, market concentration, and firm organization. Moreover, the model quantitatively matches quasi-experimental evidence on the pass-through of idiosyncratic demand shocks to wages in Portugal (Garin and Silvério, 2024), as well as the employment and wage effects of minimum wage reforms in a comprehensive set of developed economies (Dube and Zipperer, 2024).

4.1 Cross-sectional evidence

Moments of employment and wages. We begin by analyzing the wage and employment distributions, which are relevant to understanding the degree of market concentration and

Figure 1: Distribution of employment and wages across firms



(a) Employment (b) Wages

Note: The Figures show the distribution of employment and wages across firms in the model and data. In particular, the graphs plot the mean outcome for each quintile of the variable of interest.

the effects of minimum wage policies. Figure 1 displays the average level of employment and wages across quintiles in the model and data. The model closely matches that the vast majority of employees are concentrated in the top employment quintile of firms. It also captures that approximately half of firms pay wages near the minimum wage. Moreover, Panel A in Table 2 shows that the model gets right that about 90 percent of minimum wage earners are production workers. Hence, the model delivers realistic patterns regarding *how* likely and *who* is most likely to be affected by minimum wage reforms.

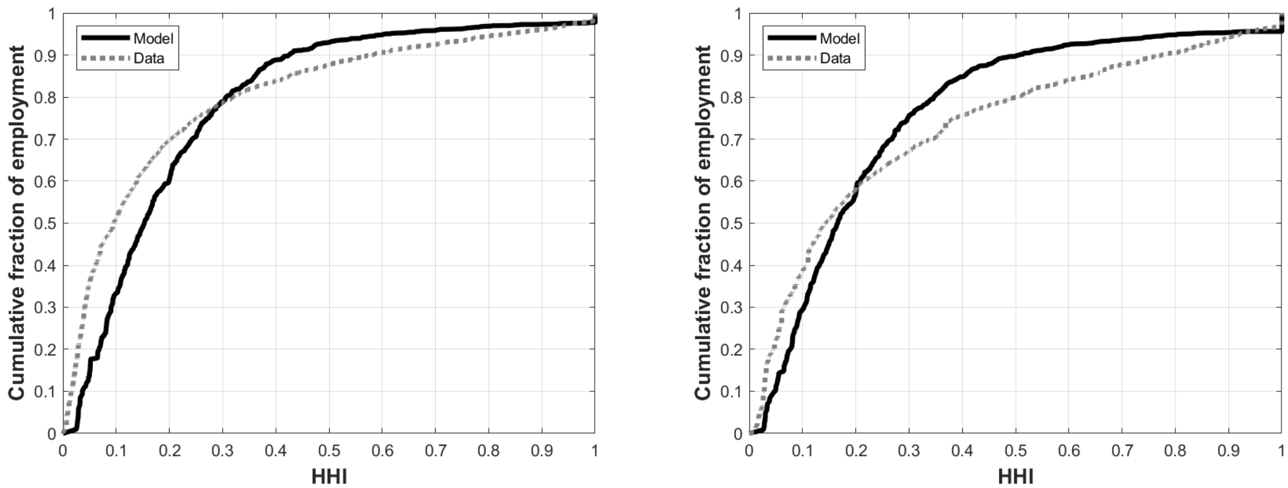
Moments of firm organization and market concentration. The distribution of workers and managers across firms is key to understanding the misallocation of labor that results from high-productivity firms exerting stronger monopsony. Panel B in Table 2 shows that the model predictions of the distribution of production workers across firms aligns with the data. Most firms are small and hire less than two employees, whereas only firms in the top one percent of the distribution hire more than 50 production workers. The model somewhat underestimates the number of managers in the top one percent of two-layer firms. However, it closely matches the data up to the 90th percentile and replicates a median span of control

Table 2: Model fit of untargeted moments

	Production Workers		**Managers**	
	Model	**Data**	**Model**	**Data**
Panel A: Minimum Wage				
Share | Minimum wage earner	0.85	0.94	0.15	0.06
Panel B: Firm Organization				
Median span of control	3.57	3.14		
P25 firm size	1	1	0	0
P50 firm size	2	2	1	1
P90 firm size	13	9	4	5
P99 firm size	55	59	9	34
Panel D: Market Concentration				
Weighted mean HHI			0.24	0.27
Weighted mean Max s_{ij}	0.31	0.30	0.34	0.38

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

Figure 2: Labor market concentration



(a) Production workers (b) Managers

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

of 3 production workers per manager among two-layer firms.

Moments of market concentration. Next, consider moments of labor market concentration, which directly speak to the level of wage markdowns and endogenously arise from agents' labor demand and supply decisions. Figure 2 shows that the model delivers a realistic pattern of the distribution of employment across markets ranked by their HHI. The model rationalizes two key features: (i) most workers sort into relatively low-concentrated labor markets, and (ii) managers tend to work in more concentrated markets than production workers. The first pattern arises because low-concentrated markets have more firms and these firms pay higher wages. The second reflects that managerial markets are more likely to exhibit high concentration levels and that managers face higher across-market mobility costs, discouraging them from leaving markets that pay relatively low wages. For the above-mentioned two reasons, the model successfully predicts average differences in market concentration across occupation. Panel C in Table 2 shows that the model closely matches that the average manager works in a market with an HHI of around one-fourth, which is about 5 percentage points higher than that of production workers. In addition, the last row in Panel C of Table 2 shows that the model captures that the average firm size in terms of payroll shares is higher for managers than production workers.

4.2 Quasi-experimental evidence

Pass-through. The pass-through of idiosyncratic demand shocks is informative about the degree of oligopsony power in the economy. Price-taking firms expand employment while taking the same wage as given in response to an idiosyncratic demand shock, so the pass-through is zero. Conversely, pass-through converges to one as firms confront increasingly inelastic labor supply curves. To validate our model in this dimension, we replicate the quasi-experiment from Garin and Silvério (2024).

GS exploit idiosyncratic export shocks during the onset of the Great Recession (2008-9) in Portugal to estimate wage pass-through by comparing wage growth across exporting firms with different exposure to unexpected shifts in foreign demand. We replicate their quasi-experiment through the following procedure. We solve the model in general equilibrium, then draw a random sample of firms and increase their idiosyncratic productivity by $\omega^{GS} \sim$ Lognormal$(\mu_\omega, \sigma_\omega)$. Moreover, we limit our sample to firms with more than $\underline{n}^{GS}$ employees. We take $\sigma_\omega = 0.05$ and calibrate the parameters $\{\underline{n}^{GS}, \mu_\omega\}$ to match an average firm size of 27 employees and an average decline of 19.4 percentage points in log value-added among affected firms, respectively.[7]

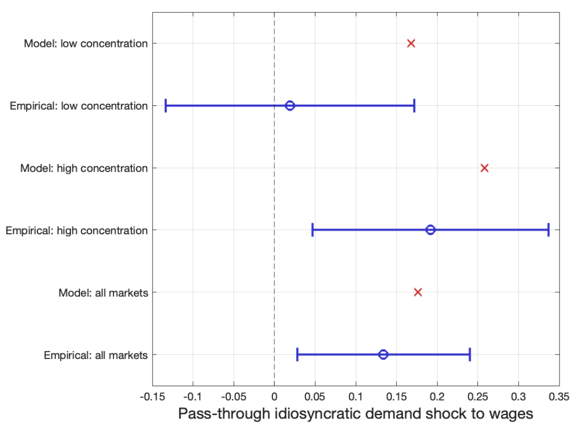
Next, we solve the model in the new general equilibrium and treat the observations before and after the shock as a panel with two time periods. We then regress log wage growth on log value added growth and measure pass-through as the regression coefficient. The left panel of Figure 3 presents the results from our model in comparison to the empirical findings. GS document a pass-through of idiosyncratic shocks to wages of 0.13.[8]. In our simulated data, we estimate a wage pass-through of 0.17, which is remarkably close and not statistically different from the GS estimate. Furthermore, GS also provides wage pass-through estimates

[7] We take the average employment size of firms from the first row of Table 2 and the average drop in value-added from the second column in Table 5. Results are robust to choosing $\sigma_\omega \in \{0.03, 0.08\}$.

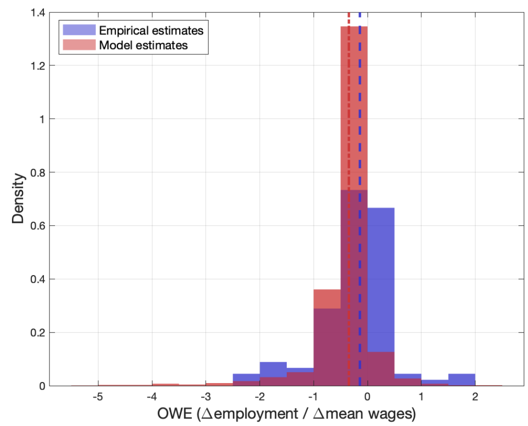
[8] This is the result from the first column of Table 7. We take monthly wages as baseline because we use it in the quantification of the model and does not significantly differ from results using hourly wages.

for the sub-sample of firms located in markets above and below the national median of the HHI. Consistent with an oligopsonistic framework, they find that wage pass-through is higher in highly concentrated markets (p-value=0.10). Using the simulated sample, we show that our model quantitatively replicates these results. Our pass-through estimates are not statistically different from GS and reproduce the higher pass-through observed in more concentrated markets.

Figure 3: Quasi-experimental evidence and model simulations



(a) Wage pass-through



(b) Effects of minimum wage policies

Note: The Figure plots the pass-through of idiosyncratic demand shocks to wages (left) and the own wage elasticities (right) from the model simulations and quasi-experimental evidence from Garin and Silvério (2024) and Dube and Zipperer (2024). We exclude the 1st and 99th percentile in the graphs for visual illustration, but all the highlighted results still hold.

Labor market effects of minimum wages. Dube and Zipperer (2024; henceforth DZ) collects a comprehensive set of estimates of the own-wage elasticity (OWE) from about 90 studies that cover a set of developed countries since 1992. The OWE measures the percent change in employment for a given percent change in the average wage induced by a minimum wage change.[9] The OWE is a meaningful measure to analyze the labor market effects of minimum wage changes because, as long as the effect of minimum wages on the mean wage is positive, an OWE > -1 implies that minimum wage reforms increase economy-wide pre-

[9]Most of the studies use data from the U.S., and the remaining studies are based on the United Kingdom, Canada, or countries in the European Union. Formally, OWE $= \frac{(\%\Delta\text{Employment})}{(\%\Delta\text{Minimum wage})} \Big/ \frac{(\%\Delta\text{Mean wage})}{(\%\Delta\text{Minimum wage})}$.

tax earnings.

The key restrictions to appear in the DZ sample require that the studies evaluate the employment effects in the statutory minimum wage and include experimental or quasi-experimental variation. To replicate this experimental evidence, we simulate a general equilibrium for all statutory minimum wages within the range $\underline{w} \in [0 : 20 : 950]$. This yields a matrix 48×48 of OWEs derived from pairwise comparisons across these economies, which only differ in terms of the statutory minimum wage.

The right panel of Figure 3 shows the histogram of empirical and model-simulated estimates of the OWE. The dashed lines in the Figure display the mean value of each distribution. Two results show that the model quantitatively replicates the labor market effects of minimum wage reforms. First, the model generates most of the OWEs within within the range of empirical estimates. Second, mean OWE is -0.2 in DZ, while the model-simulated mean is -0.4. The model somewhat overestimates the negative employment effects of minimum wages, but correctly captures that, on average, raising the statutory minimum wage leads to fairly modest employment losses and increases pre-tax earnings.

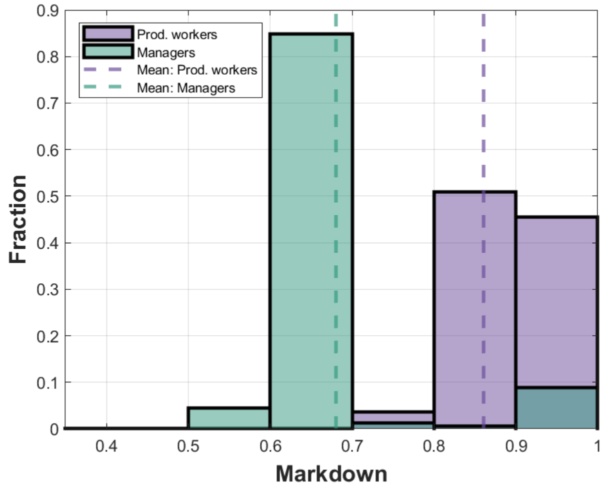
5 Implications of Occupation-Specific Monopsony Power

Firms' monopsony over managers is key to understanding the overall efficiency and welfare losses from monopsony power. The average markdown is twice as high for managers (32 percent) as for production workers (16 percent). Relative to an efficient economy, these wedges induce a Utilitarian social welfare loss of 2.7 percent. The direct effect of managers' markdowns, along with spillovers arising from production complementarities between occupations, account for 0.9 pp. of this loss. While managers always enjoy welfare gains within a reasonable profit-share range, production workers' welfare only rises when they bear sufficiently low profit losses.

5.1 Measuring Monopsony Power

Figure 4 displays the distribution of wage markdowns for both occupations in the benchmark economy. Markdowns are below one due to imperfect firm substitutability and firm granu-

Figure 4: Distribution of wage markdowns



Note: The Figure plots the distribution of wage markdowns (μ_{ij}) across firms for production workers and managers. 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.

larity, implying that wages are below the marginal revenue product of labor. We estimate an employment-weighted markdown of 16.0 percent for production workers and 31.9 percent for managers. Three reasons explain why monopsony is stronger over managers than production workers. First, both the upper (η_o) and lower bounds (θ_o) of the structural elasticities are lower for managers. Namely, compared to production workers, managers perceive distinct firms as more imperfect substitutes in terms of non-wage characteristics. Second, firms tend to have higher managers' than production workers' payroll shares. Thus, the labor supply elasticity of managers is closer to the across-market elasticity than that of production workers. In other words, managers find it harder to reallocate toward other firms because more of their alternatives are outside of their current local market. Third, minimum wages mainly constrain low-productivity firms, which primarily employ production workers.

5.2 Welfare and Efficiency Losses from Monopsony Power

The presence of wage markdowns translates into pure deadweight and misallocation losses. Thus, we now study the effect of monopsony power on efficiency and welfare. We compute the

Table 3: Effects of efficient economy relative to benchmark with monopsony

	Efficient economy (% change)	Manager effects (pp.)		Efficient economy (% change)	Manager effects (pp.)
Panel A: Mean wages					
Production workers	20.5	4.6	Managers	46.8	43.2
Panel B: Employment					
Production workers	7.9	1.7	Managers	14.6	14.5
Total	9.1	4.1	Output	10.1	5.9
Panel C: Firm organization					
Share two-layer firms	-10.9	-4.5	Span of control	-4.4	-0.4
Panel D: Mean HHI					
Production workers	20.0	4.6	Managers	20.4	14.6

Note: The Table reports the percent change in outcomes in the efficient (markdowns equal to one) relative to the benchmark economy with monopsony power. In addition, the Table also shows the contribution of managers' monopsony power to the overall effects. We compute the latter as the difference between (i) the efficient economy and (ii) a counterfactual with markdowns equal to one exclusively for production workers.

efficient allocation by setting wage markdowns to one, i.e., equalizing wages to the marginal product of labor for all firms. Note that this counterfactual does not change households' preferences. Table 3 summarizes the results by comparing the aggregate outcomes in the efficient relative to the benchmark economy.

Panel A in Table 3 shows the change in mean wages across occupations. Since managers bear wider markdowns, the mean wage increase is greater for managers than for production workers. As a result, wage dispersion across occupation increases. Panel B shows the employment effects of monopsony power. When firms face upward sloping labor supply curves, they restrict employment relative to the efficient allocation to reduce labor costs. We find that employment rises by 7.9 and 14.6 percent for production workers and managers in an efficient economy, respectively.

In an efficient economy, wages and employment increase, as well as the concentration of employment at the most productive firms. The top panels in Figure 5 show that the share of employees, especially managers, working in the most productive firms rises by more than 15 percent. Moreover, the bottom left panel shows that part of this increase in concentration stems from the fact that rising wages make management delegation no longer profitable for medium-productivity firms, with the share of two-layer firms decreasing by 10.9 percent. Overall, employment gains and reallocation together increase output by 10.1 percent.

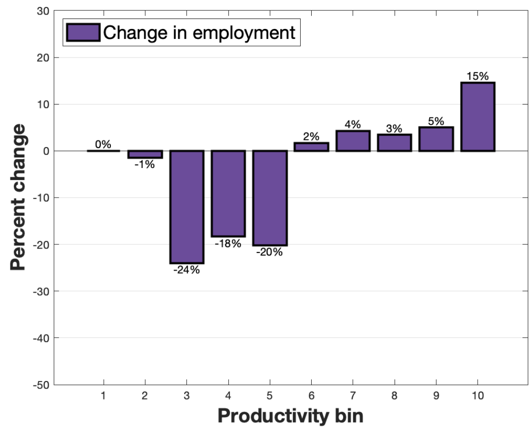
The large and heterogeneous wedges between marginal productivity and wages become a major detriment to households' welfare. We use a utilitarian welfare function that defines social welfare as the sum of utilities of both occupations: $\mathcal{U}_S = \sum_{o \in \{w,m\}} \mathcal{U}_o$. Note that the size of the household enters the utility function by affecting aggregate consumption and labor supply. We find a consumption equivalent social welfare gain of 2.7 percent, which stems from consumption gains and despite worker reallocation and the profit losses.

Linear utility in consumption implies that the distribution of firm profits across occupations, κ_o, does not affect social welfare. In contrast, Figure 5(d) shows that the share of profits rebated to a specific household significantly affects the impact of monopsony power on each household's welfare. We consider a reasonable range for the share of profits rebated to production workers to be between equal shares and population shares, $\kappa_w \in [0.4, 0.8]$. Within this interval, the welfare change of production workers between the efficient and benchmark economies lies in the range $\xi_w \in [-0.02, 0.04]$. When the share is slightly below equal shares ($\kappa_w = 0.4$), production workers enjoy a welfare gain of up to 4.3 percent. In contrast, when they suffer the bulk of the profit losses ($\kappa_w = 0.8$), these losses largely offset the relatively small gains from higher earnings, and they experience a welfare loss of 2.5 percent. Unlike production workers, managers always enjoy welfare gains over the same range of profit-sharing allocations, $\xi_m \in [0, 0.18]$, reflecting the strong monopsony power firms hold over them, with welfare gains of up to 18.1 percent.

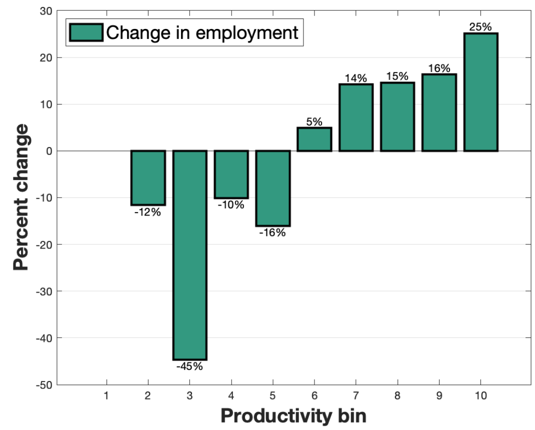
5.3 The Role of Managers' Monopsony

In our framework, where production complementarities exist between occupations, the joint distribution of allocations and prices across occupations is mutually dependent. Hence, monopsony power over one occupation spills over to others and has far-reaching consequences for the entire workforce. To isolate this channel, we simulate a counterfactual economy where

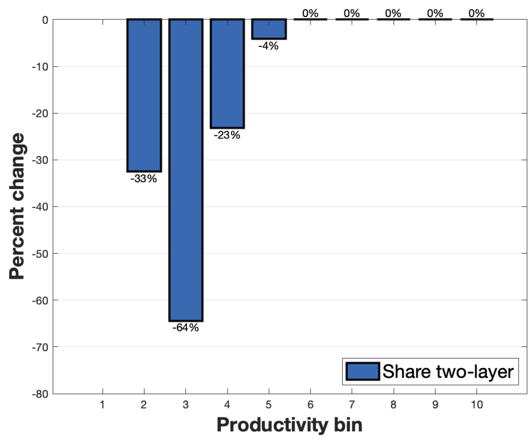
Figure 5: Effect of monopsony power on misallocation and welfare



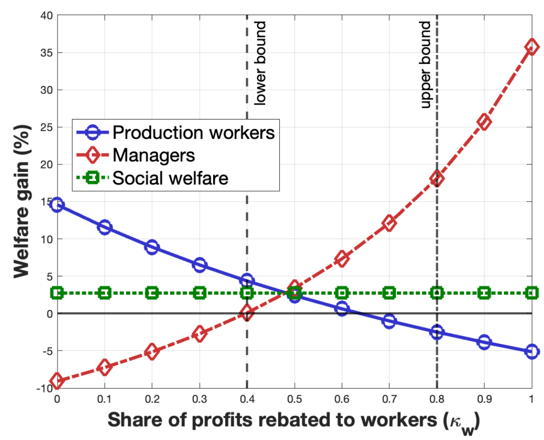
(a) Production workers



(b) Managers



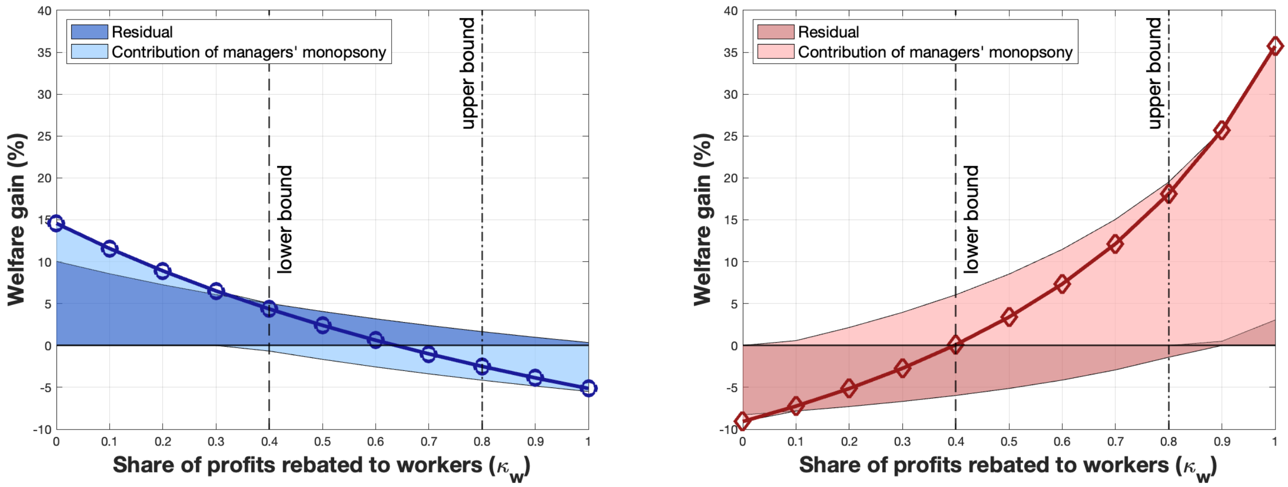
(c) Share two-layer firms



(d) Welfare

Note: The top panels in the Figure plot the percent change in employment of production workers (left) and managers (right) across firms in the efficient relative to the benchmark economy. The bottom left panel shows the percent change in the share of two-layer firms, where we express the change as a fraction over the total number of firms in the same productivity bin. The bottom right panel shows the social and occupation-specific welfare gains in (percent) consumption equivalent terms. The efficient economy consists of an economy where wages are equal to the marginal product of labor. We classify firms into ten bins according to their productivity, where a higher bin implies higher productivity.

Figure 6: The role of managers' monopsony in welfare



(a) Production workers (b) Managers

Note: The Figures display how much the firm organization channel, i.e., the endogenous firms' choice of layers, accounts for the percent change between the efficient and benchmark economies in several outcomes of production workers (left) and managers (right). For instance, the endogenous organizational choice of firms explains about 30 percent of the change in the average level of payroll concentration in managerial markets between the efficient and benchmark economies.

we exogenously set wage markdowns to one exclusively for production workers. We then attribute the differences between this counterfactual and the efficient economy to managers' monopsony and its interactions with production workers' monopsony. We refer to both the *direct* and *spillover* effects as the *effects from managers' monopsony* throughout the sequel.

Table 3 shows the contribution, in percentage points, of managers' monopsony to the overall changes between the efficient and benchmark economies. Moving to an economy with efficient managers' wages explains almost entirely the changes in their own labor market outcomes and a significant part of the changes in production workers' outcomes. Rising managerial wages, especially at the most productive firms, results in an increase in both overall employment and concentration of managers at the most productive firms. Combined with production complementarities, rising managers' employment incentivizes these firms to expand their hiring of production workers and their production. Specifically, nearly one-fifth of the increase in market concentration (4.6 pp.), employment (1.7 pp.), and wages (4.6 pp.)

of production workers solely stems from managers' monopsony. Moreover, efficient managers' wages explain about 60 percent of the output gains in the efficient economy.

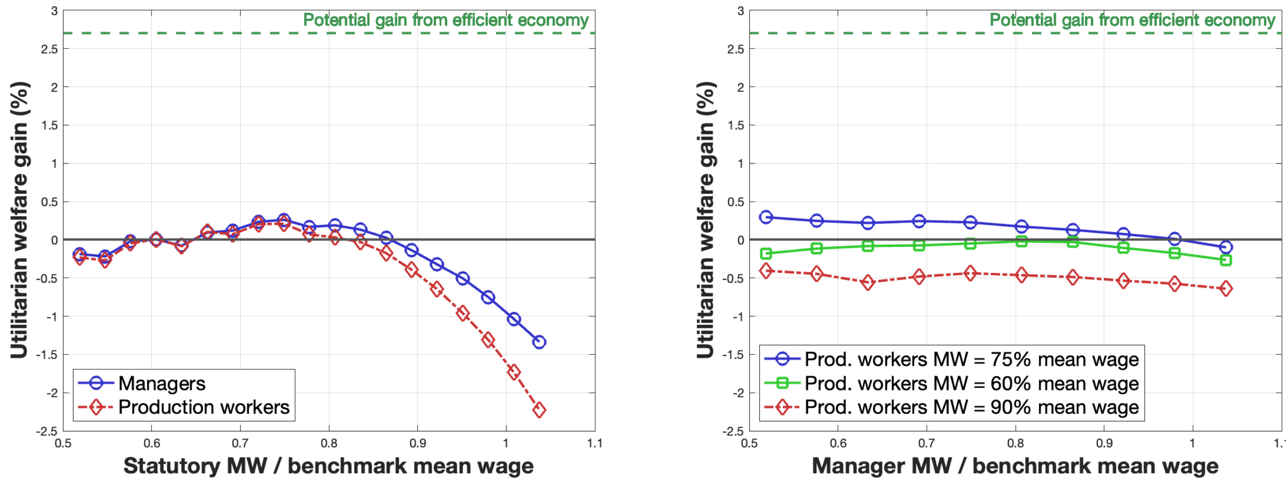
Regarding welfare, we find that removing managers' monopsony increases social welfare by 0.9 pp., representing about one-third of the gains from the efficient economy. Figure 6 shows the contribution of managers' monopsony to each occupation's welfare changes from the efficient economy, with the remaining share (residual) attributed exclusively to production workers' monopsony. For each occupation, moving to an economy with efficient wages in its own occupation brings about welfare gains for all profit shares, κ_o, due to the large earnings gains. However, workers in one occupation does not necessarily benefit from removing monopsony power over the other occupation. While their earnings rise as firms expand employment and wages due to production complementarities, this effect is counteracted by a decline in firm profits and the reallocation of workers. The positive effect only outweighs the negative one when the profit share of the occupation is relatively low.

The above-mentioned results emphasize that policies targeting monopsony power exclusively in low-wage occupations would prove ineffective at addressing substantial welfare losses that also stem from monopsony power over high-wage occupations. This finding motivates our assessment of minimum wage policies' effectiveness in capturing the efficiency and welfare gains from an efficient economy.

6 Minimum Wage Policies

Among other reasons, minimum wage policies aim to improve the well-being of low-income workers by reducing the wage-setting power of firms. We find that the single statutory minimum wage that maximizes social welfare captures less than 10 percent of the social welfare gains from an efficient economy. Despite the presence of occupation-specific monopsony power, an optimal occupation-based minimum wage only slightly improves upon the optimal single minimum wage. The reason is that, conditional on an occupation, minimum wages inevitably bind first for low-productivity firms where monopsony power is relatively weak.

Figure 7: Welfare effects of minimum wage policies



(a) Statutory minimum wage

(b) Occupation-based minimum wage

Note: The Figure plots the occupation-specific (left) and social (right) consumption equivalent gains in a counterfactual relative to the benchmark economy. On the left panel, the only difference between the counterfactual and benchmark economy is a different statutory minimum wage. On the right panel, the counterfactual simulations differ in terms of occupation-specific minimum wages.

6.1 The Optimal Statutory Minimum Wage

The Portuguese government has significantly increased the (real) statutory minimum wage during the last two decades, with the share of employees earning the minimum wage reaching one-fourth by 2017 (see Figure A.7). In the model, raising the minimum wage mitigates monopsony power in a firm by inducing this firm to increase employment and wages as long as there is a wedge between wages and the marginal revenue product of labor.

Figure 7(a) plots the welfare effects of different statutory minimum wages relative to the benchmark economy. Welfare is hump-shaped and attains the same welfare maximizing minimum wage for both occupations at a minimum to mean wage of about 75%. At best, an optimal statutory minimum wage captures less than one-tenth (0.2 pp.) of the social welfare gains from an efficient economy (2.7 percent). Despite generally large markdowns, the statutory minimum wage fails to effectively address monopsony due to the presence of firm and occupational markdown heterogeneity. Conditional on occupation, the statutory minimum wage binds first for low-productivity firms, which exert relatively narrow markdowns. Similarly, conditional on firm productivity, the statutory minimum wage is more likely to bind for low-wage occupations, which also display relatively smaller markdowns. As a result, a statutory minimum wage is not able to address the strong monopsony over high-wage workers without large losses in output, employment, and ultimately, welfare.

6.2 The Optimal Occupation-Based Minimum Wage

The occupational heterogeneity in wage markdowns suggests that designing a minimum wage for each occupation, rather than a single statutory minimum wage, could be more effective at tackling the welfare losses from monopsony power, as managers tend to earn wages above the minimum wage and bear wider markdowns than production workers. Indeed, many developed countries already implement occupations-based wage floors.[10]

For the above-mentioned reasons, we simulate different scenarios where we set a specific minimum wage for each occupation. Figure 7(b) depicts the change in the Utilitarian social welfare for different values of occupation-specific minimum wages. Each line represents the social welfare gains for varying manager-specific minimum wages, conditional on a fixed production worker minimum wage. For clarity, we display only the (blue) line representing the welfare-maximizing combination of minimum wages, along with two additional lines.

We find that the combination of occupation-based minimum wages that maximizes social welfare provides a gain of 0.3 percent relative to the benchmark. This occurs when the minimum wage of production workers is about 75 percent of their mean wage and the minimum wage of managers is about 50 percent of their mean wage. Two things stand out from this result. First, despite strong monopsony power over managers, a social planner would set their minimum wage relatively low. This is because social welfare mostly depends on

[10] Under the Modern Awards system, the Australian government implements statutory minimum wages that are occupation-based. Moreover, collective agreements that set distinct wage floors across occupations in an industry are common in European countries such as Italy (Adamopoulou et al., 2023), France (Fougère et al., 2018), Belgium, Sweden (International Labour Organization, 2023), and Portugal. However, Portuguese workers typically earn significant wage premiums over their wage floor (Card and Cardoso, 2022), indicating that wage bargaining is not effective in preventing firm-wage setting power for most employees.

production workers' outcomes, and the managerial job losses from a higher manager-specific minimum wage reduce labor demand and welfare for production workers. Second, optimal occupation-specific minimum wages provide only marginal improvement over the optimal single statutory minimum wage. Like the single minimum wage, they capture relatively modest welfare gains compared to an efficient economy, as they primarily affect low-productivity firms where monopsony power is weaker.

7 Conclusion

Extending a general equilibrium model of oligopsony to include minimum wages and firm organization, we show that the heterogeneity in monopsony power and production complementarities across worker types helps understanding the firm size distortions and welfare effects of monopsony power. We estimate the model to employer-employee-matched data and balance sheet data from Portugal and validate it against quasi-experimental evidence. We quantify an average wage markdown of 31.9 and 16.0 percent over managers and production workers, respectively. Consequently, moving from the benchmark to an efficient economy increases employment, mean wages, and employment concentration at the most productive firms, especially for managers. As a result of the production complementarities, nearly one-fifth of these gains for production workers stem from eliminating managers' monopsony power alone. Overall, the efficient economy provides a consumption-equivalent social welfare gain of 2.7 percent relative to the benchmark, as the higher earnings offset the profit losses and utility costs from worker reallocation. Under equal profit shares, managers and production workers enjoy a welfare gain of 3.4 and 2.4 percent, respectively.

We consider two valuable extensions for future research. We believe market concentration and mobility costs may also show a systematic relationship with other market characteristics. For instance, the skill level required in the job, workers' age, or whether the market is formal or not in developing countries. Regarding the model specification, future model extensions may consider worker heterogeneity in productivity and across-occupation mobility.

References

Adamopoulou, E., F. Manaresi, O. Rachedi, and E. Yurdagul (2023): "Minimum Wages and Insurance within the Firm," .

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

Azkarate-Askasua, M. and M. Zerecero (2024): "Union and firm labor market power," *Available at SSRN 4323492*.

Bachmann, R., G. Demir, and H. Frings (2022): "Labor market polarization, job tasks, and monopsony power," *Journal of Human Resources*, 57, S11–S49.

Bamford, I. (2021): "Monopsony Power, Spatial Equilibrium, and Minimum Wages," *Unpublished manuscript*.

Barkai, S. (2020): "Declining labor and capital shares," *The Journal of Finance*, 75, 2421–2463.

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

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

——— (2025): "Minimum Wages, Efficiency, and Welfare," *Econometrica*, 93, 265–301.

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

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

Caliendo, L. and E. Rossi-Hansberg (2012): "The impact of trade on organization and productivity," *The Quarterly Journal of Economics*, 127, 1393–1467.

Card, D. and A. R. Cardoso (2022): "Wage flexibility under sectoral bargaining," *Journal of the European Economic Association*, 20, 2013–2061.

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

Deb, S., J. Eeckhout, A. Patel, and L. Warren (2024): "Walras–Bowley Lecture: Market power and wage inequality," *Econometrica*, 92, 603–636.

Diamond, R. (2016): "The determinants and welfare implications of US workers' diverging location choices by skill: 1980–2000," *American Economic Review*, 106, 479–524.

Drechsel-Grau, M. (2023): "Employment and Reallocation Effects of Higher Minimum Wages," *Working Paper*.

Dube, A. and B. Zipperer (2024): "Own-wage elasticity: Quantifying the impact of minimum wages on employment," Tech. rep., National Bureau of Economic Research.

Fougère, D., E. Gautier, and S. Roux (2018): "Wage floor rigidity in industry-level agreements: Evidence from France," *Labour Economics*, 55, 72–97.

Garcia-Santana, M. and J. Pijoan-Mas (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, 874–904.

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

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

Garin, A. and F. Silvério (2024): "How responsive are wages to firm-specific changes in labour demand? Evidence from idiosyncratic export demand shocks," *Review of Economic Studies*, 91, 1671–1710.

Goolsbee, A. and C. Syverson (2023): "Monopsony power in higher education: A tale of two tracks," *Journal of Labor Economics*, 41, S257–S290.

Grobovsek, J. (2020): "Managerial Delegation, Law Enforcement, and Aggregate Productivity," *The Review of Economic Studies*, 87, 2256–2289.

Heise, S. and T. Porzio (2023): "Labor Misallocation Across Firms and Regions," Tech. rep.

Hsieh, C.-T. and P. J. Klenow (2009): "Misallocation and manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 124, 1403–1448.

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

International Labour Organization (2023): "A Review of Wage Setting through Collective Bargaining," Tech. rep., Geneva: International Labour Office.

Jarosch, G., J. S. Nimczik, and I. Sorkin (2023): "Granular Search, Market Structure, and Wages," *Working Paper*.

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

Karabarbounis, L., J. Lise, and A. Nath (2022): "Minimum Wages and Labor Markets in the Twin Cities," Tech. rep., National Bureau of Economic Research.

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

Lamadon, T., M. Mogstad, and B. Setzler (2022): "Imperfect competition, compensating differentials, and rent sharing in the US labor market," *American Economic Review*, 112, 169–212.

Langella, M. and A. Manning (2021): "Marshall Lecture 2020: The Measure of Monopsony," *Journal of the European Economic Association*, 19, 2929–2957.

Lawson, N., C. Lelarge, and G. Spanos (2023): "The Minimum Wage in Firms' Organizations: Productivity Implications," .

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

MacKenzie, G. (2021): "Trade and market power in product and labor markets," Tech. rep., 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*.

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

Santamaria, C. (2023): "Small Teams in Big Cities: Are Technology Differences Across Cities Optimal?" *Working Paper*.

Severen, C. (2021): "Commuting, labor, and housing market effects of mass transportation: Welfare and identification," *Review of Economics and Statistics*, 1–99.

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

Tamkoç, N. M. (2022): "Production Complexity, Talent Misallocation and Development," *Working Paper*.