Job Preferences, Labor Market Power, and Inequality

Oscar Volpe

October 31, 2024

[Access the **latest draft** and **online appendix**.]

Abstract

I analyze how a firm's labor market power shapes, and is shaped by, its workforce, and I evaluate the implications for welfare and inequality. Using matched worker-firm panel data from Norway (1995-2018), I develop, identify, and estimate an equilibrium model of the labor market where firms compete with one another for workers who are heterogeneous in both their skills and preferences over wages versus non-wage job amenities. I allow the wage-amenity trade-offs to be correlated with skills, while also varying among equally skilled workers. When a firm adjusts its wages, the composition of its workforce shifts, and these compositional changes, in turn, affect the labor supply curve to the firm. As a result, the firm's wage-setting power varies based on which types of workers it employs. I find that this variation leads to large allocative inefficiency, with welfare losses from imperfect competition estimated at 9.5% relative to the competitive benchmark.

I. Introduction

The role of firms in shaping wage inequality is an important topic for labor policy. Three growing empirical literatures examine how firm characteristics influence workers' earnings. One body of research (Manning, 2021; Card, 2022; Azar & Marinescu, 2024) studies the role of imperfect competition in the labor market, where firms exert wage-setting power due to facing upward-sloping labor supply curves. A second literature—building on work by Abowd, Kramarz, & Margolis (1999)—focuses on sorting, using matched employer-employee data to uncover the relationship between worker and firm productivity. A third literature examines

I am grateful to my advisors Alex Torgovitsky, Magne Mogstad, Steven Durlauf, and Thibaut Lamadon for their invaluable guidance and encouragement. I also thank Alan Manning, Evan Rose, Derek Neal, Manasi Deshpande, Michael Dinerstein, Michael Greenstone, Mikhail Golosov, Stéphane Bonhomme for detailed and insightful comments, as well as Arnstein Vestre, Jeanne Sorin, Hugo Lopez, Lancelot H. de Frahan, Michael Graber, Ola Vestad, Tanya Rajan, Tom Melling, and the participants at the Applied Micro working groups and Public/Labor lunches at the University of Chicago for helpful feedback and suggestions. All errors are my own.

*Department of Economics, University of Chicago.

workers' willingness to trade higher wages for better non-wage aspects of a job, finding that the trade-offs vary substantially among workers (Mas & Pallais, 2017; Wiswall & Zafar, 2018).

While often studied separately, the issues explored by these literatures are interconnected. In a labor market with differentiated employers, a firm's wages and non-wage amenities determine which workers it attracts. Furthermore, if workers make different trade-offs between wages and amenities, then the composition of a firm's workforce could matter for the elasticity of labor supply that it faces, which in turn impacts the firm's wage-setting power. These factors prompt a key question: *How is a firm's wage-setting power linked to its workforce?* Addressing this question would offer deeper insight into how workers sort and how market power varies within the labor market, which has important implications for welfare and inequality.

This paper develops an empirical framework to examine how workers choose jobs based on heterogeneous wage-amenity trade-offs, and the consequences of this behavior for a firm's wage-setting power. I show that the market power at a firm—as measured by wage markdowns and rents—directly depends on the composition of its workforce, which is influenced by the wages and amenities the firm provides. Moreover, by failing to account for this relationship, one would overlook a key source of allocative inefficiency that arises from differences in wage markdowns among heterogeneous firms. I show how to recover economic quantities under my framework using matched worker-firm panel data, while allowing for unobserved heterogeneity in worker skills, firm productivity, technology, and amenities. I then apply my method to administrative data from Norway, where I draw inference about the determinants of worker sorting, the concentration of market power, and the welfare impacts of imperfect competition.

I begin my analysis by presenting an equilibrium model of the labor market. Building on previous models of job differentiation, such as Card et al. (2018), Lamadon et al. (2022), and Azar et al. (2022), I assume that workers' idiosyncratic tastes are not fully priced into wages, leading firms to face imperfectly elastic labor supply curves. However, unlike previous work, I allow workers to have varying marginal rates of substitution between wages and amenities. Drawing on the literature about the willingness to pay for amenities, I allow the wage-amenity trade-offs to be correlated with skills, while also differing among similarly skilled workers.

This extension has key implications for the shape of the labor supply curve faced by a firm. It establishes a link between worker and firm characteristics, such that the elasticity of labor to the firm depends on the composition of its workforce. Consequently, the elasticities differ among firms, and they are also endogenous to changes in the environment: if a firm adjusts its wages, then the composition of its workforce would shift, and these compositional changes would, in turn, influence the elasticities. This variation contrasts with the standard assumption of homogenous and isoelastic labor supply curves, which leads to uniform wage markdowns. Moreover, to the extent that workers' wage-amenity trade-offs are correlated with their skills, wage markdowns may also differ within a firm—among workers with varying skills—as well as across firms. These implications do not rely on firms competing strategically for workers, as in Berger et al. (2022). Even when firms are strategically small within the labor market, their wage-setting power would still vary based on *who* is employed at each firm.

To account for many potential sources of worker sorting, I allow firms to be differentiated along several key dimensions. Specifically, I allow firms to vary in productivity, such that a worker's skills may be valued differently at different firms. I also allow for heterogeneous technology in the production function, where firms can exhibit decreasing returns to scale and imperfect substitutability between labor inputs. Moreover, I assume that firms have distinct non-wage amenities, which also differ within a firm based on workers' skills.[1] I demonstrate how each of these firm characteristics interacts with workers' preferences in equilibrium, and I discuss the consequences of these interactions for wage markdowns, rents, and misallocation.

I then turn to the question of identification, demonstrating how to learn about imperfect competition under my model using matched worker-firm data. In my analysis, identification is challenging for two reasons. First, as is typical in most contexts, a worker's wage markdown is not observed in the data. In particular, observationally similar workers could earn different wages due to unobserved skill differences—not market power.[2] Wages may also not fully reflect a worker's compensation, as firms may have amenities that researchers do not observe. Indeed, as Rosen (1986) discusses, compensating differentials can exist even in competitive markets. A second identification challenge stems from the generality of my framework, which allows workers' wage markdowns to vary both within and across firms, and to be endogenous to wage changes within a firm. This added feature complicates my analysis because instead of recovering a single markdown, I need to recover entire functions of firm characteristics.

I show how to overcome these challenges to achieve identification of the model through the use of panel data by leveraging an economic shock that affects the productivity of multiple firms in the economy. These productivity shocks can be recovered either from existing data using internal instruments or from external data sources. One novelty of my approach is that I allow the productivity shocks to shift wages in a way that is reflected in workers' wage indices. Therefore, I allow there to be spillover effects on the labor supply curves faced by *all* firms, including those that do not experience a productivity shock. These spillovers would prevent me from recovering the elasticity of labor supply to a firm by directly comparing a firm's outcomes before and after the shocks. Nevertheless, I prove that—under weak conditions on

[1]Recent empirical work provides mixed evidence about which non-wage job characteristics are most important to workers; for discussion, see Maestas et al. (2017). In my analysis, I do not take a stance on this issue. Instead, I assume that firms are endowed with a fixed set of amenities that are unobserved by the analyst; or, more precisely, I restrict amenities to be fixed over the estimation window. This assumption neither imposes nor precludes the possibility that firms initially choose amenities to maximize profits. Indeed, it is straightforward to show that permitting firms to choose their amenities initially would have no impact on any of my estimates.

[2]See, e.g., Murphy & Topel (1990), Gibbons & Katz (1992), Abowd et al. (1999), and Gibbons et al. (2005).

the instrument—the elasticities are identified from a difference-in-differences comparison of outcomes. This method involves comparing mean changes in wages and labor between firms with different exposures to productivity shocks, while controlling for a firm's initial wages.

An additional contribution of my analysis is that I draw inference about the differences in workers' trade-offs between wages and amenities. This approach is based on a revealed preference argument: holding all else fixed, a firm that offers better wages (for a given skill group) is more likely to attract workers who value wages more highly. This compositional effect is captured by the labor supply elasticity faced by a firm. Given this property, I prove that there exists a one-to-one mapping between the density function governing workers' wage-amenity trade-offs and a firm's labor supply curve. I then develop a method to estimate this density nonparametrically, which can be implemented separately for different population subgroups.

Using this identification strategy, I demonstrate how to recover the structural parameters in the model that characterize firm productivity, technology, and amenities, as well as workers' preferences. From these parameters, I can calculate the rents for both workers and firms, and conduct counterfactual analyses to explore sources of allocative inefficiency and inequality.

Finally, I estimate the model using a matched employer-employee panel dataset, covering the universe of workers and firms in Norway for the period 1995-2018. My results yield three key findings. First, I find that many firms have substantial wage-setting power. On average, workers' wages constitute only 86% of their marginal products. Workers and firms also earn large rents: on average, workers are willing to forgo 17% of their wages to remain at their current job, while the average firm derives 29% of its profits by exerting wage-setting power.[3]

Second, I find that labor market power varies significantly within and across firms, with the wage-to-marginal-product ratio ranging from 0.60 to 0.95. I also find that market power is more concentrated among firms that pay lower wages for a given skill group. This finding contrasts with the traditional view that larger, more productive firms should have more wage-setting power over workers.[4] However, it supports a key prediction of my model. Specifically, it suggests that lower-paying firms are more likely to employ workers who are less responsive to wage changes (less elastic), thereby enabling these firms to exert greater wage-setting power.

Third, I find that imperfect competition generates substantial misallocation of workers to firms. By eliminating labor wedges, I estimate that total welfare would increase by 9.5%, measured in dollar terms. Moreover, in a competitive (Walrasian) labor market, more workers

[3]Empirically, it is hard to credibly differentiate between ex ante and ex post firm rents. This task necessitates further information—or assumptions—about how firms choose and pay for the amenities provided to workers.

[4]This view traces back to the early research on monopsony (Robinson, 1933), which examines a labor market with a single buyer of labor. However, if labor markets contain many firms, then market power need not increase with a firm's size (Manning, 2021). Indeed, my estimates reveal that less productive firms tend to have greater wage-setting power because of they are more likely to attract workers who are less responsive to wage changes.

would be employed at firms with lower productivity. This key source of inefficiency would not be captured by frameworks that do not account for heterogeneity in workers' wage markdowns.

My paper contributes to three literatures. First, it relates to a large literature on imperfect competition in the labor market. Manning (2021), Sokolova & Sorensen (2021), Card (2022), and Azar & Marinescu (2024) provide recent reviews of this research.[5] Within this literature, my paper builds on work by Card et al. (2018), Lamadon et al. (2022), and Azar et al. (2022), where imperfect competition arises from idiosyncratic tastes for jobs that are not fully priced into workers' earnings.[6] This work, which Manning (2021) terms "new classical monopsony," uses differentiated demand models that are often employed in Industrial Organization (IO). A limitation of these approaches, however, is that they often specify utility functions where worker and firm characteristics do not interact. As a result, they place strong restrictions on the way that workers sort, which limits how firm wage-setting power varies in the economy.[7] I relax these restrictions by extending the model to allow workers to exhibit heterogeneous wage-amenity trade-offs. This extension is similar to one that is already well-established in the context of product markets, based on work developed by Berry et al. (1995). Moreover, as I demonstrate, it significantly broadens the scope of worker behavior in the labor market. In doing so, it leads to a richer characterization of the determinants of firm wage-setting power.

Second, my paper contributes to an empirical literature on sorting, drawing on the additive worker and firm effects wage model introduced by Abowd, Kramarz, & Margolis (1999). This work has largely focused on skill-based sorting, arising, e.g., from worker-firm production complementarities (Song et al., 2018). While my model allows for sorting on workers' skills, I also study a second type of sorting based on workers' wage-amenity trade-offs. I show that this feature introduces a firm-skill interaction term—the wage markdown—in a worker's wage equation, such that the same firm can have a larger markdown for one skill group and a smaller markdown for another. In doing so, my paper provides additional insight into assortativeness in the allocation of workers to firms, e.g., Andrews et al. (2008) and Card et al. (2013).

Third, my paper contributes to a growing literature on the willingness to pay for non-wage amenities. Both experimental studies and survey data offer strong evidence that workers value a wide variety of non-wage job attributes (Hamermesh, 1999; Pierce, 2001; Maestas et al., 2017). Additionally, the willingness to pay for these amenities differs among workers (Mas & Pallais, 2017; Wiswall & Zafar, 2018), contributing to wage inequality (Maestas et al., 2023). My paper takes this research area in a new direction, exploring how differences in the

[5]For additional discussion, I refer to Boal & Ransom (1997), Bhaskar et al. (2002), and Manning (2011).

[6]Similar equilibrium models are also studied by Kroft et al. (2021) and Chan et al. (2024), among others.

[7]Abstracting from strategic interactions, it leads to uniform wage markdowns. As discussed in Section II.E, this restriction can be relaxed by using a nested logit model (Lamadon et al., 2022; Azar et al., 2022) or by letting markdowns vary within a firm based on worker observables (Chan et al., 2024; Deb et al., 2024). Yet, these approaches still cannot capture sorting patterns arising from unobserved variation in wage-amenity trade-offs.

valuations of amenities affect the wage-setting power of firms in imperfectly competitive labor markets. I demonstrate that workers' wage-amenity trade-offs are reflected in the labor supply elasticities faced by firms. Moreover, by recovering these elasticities, I can draw inference about differences in wage-amenity trade-offs among workers. This method provides a new way to learn about variation in the willingness to pay for amenities using revealed preference, contributing to a recent literature in this area, e.g., Sorkin (2018) and Hall & Mueller (2018).

This paper is organized as follows. Section II presents the model and explores equilibrium properties related to sorting, imperfect competition, and welfare. Section III discusses the data sources and sample selection. Section IV outlines the identification strategy. Section V describes the estimation procedure, parameter estimates, and fit. Section VI presents the empirical findings about rents, misallocation, and inequality. Lastly, Section VII concludes.

II. Model of Imperfect Competition in the Labor Market

In this section, I develop a theoretical framework to explain how workers sort into jobs and the implications of this sorting for employer wage-setting power. My framework extends prior equilibrium models of the labor market, such as Kroft et al. (2021), Lamadon et al. (2022), and Azar et al. (2022), by allowing workers to exhibit varying marginal rates of substitution with respect to wages and non-wage job characteristics. I show that this extension leads employers to face heterogeneous labor supply elasticities, which reflect differences in the composition of workers at different firms. As a result, the model generates variable wage markdowns and rents, thus introducing a potential source of allocative inefficiency. Throughout my analysis, I allow for unobserved heterogeneity in worker skills, firm production, and non-wage amenities.

II.A. Environment

The economy comprises a unit measure of workers, indexed by i, and a finite number of employers $j \in \{1, \dots, J\}$. Each employer operates within a local labor market $m(j)$, and $\mathcal{J}_m$ denotes the set of employers in market m. In this economy, workers are differentially productive, and they exhibit heterogeneous tastes for jobs. Meanwhile, employers provide differentiated work environments, and they also vary in their productivity and technology.

Each worker i has a vector of skills $X_i = (\chi_i, \varphi_i)$. These skills contain two components: a categorical *skill type* $\chi_i \in \mathcal{X}$ and a continuous *skill level* $\varphi_i \in \mathbb{R}_+$. The worker's skill type χ_i captures the non-ordinal aspects of human capital such as training and specialization, which are valued differently at different firms. The worker's skill level φ_i, which is unobservable to the analyst, measures individual productivity within a given skill type. By characterizing skills as a composite of types and levels, the model adopts approaches from a growing literature about the importance of multidimensional skills, e.g., Lise & Postel-Vinay (2020), while also allowing human capital to vary in ways that researchers cannot directly observe.

Throughout my analysis, I maintain two assumptions about the information structure. First, I abstract from search frictions by assuming that workers are fully informed about job opportunities and that, given posted wages, a worker freely chooses where to work. Second, I assume that employers observe workers' skills $\{X_i\}_i$, but they only know the distribution of workers' preferences conditional on X_i. Therefore, while employers can assign wages based on skills, they cannot further price-discriminate with respect to individual preferences.[8]

II.B. Worker Preferences

Workers have heterogeneous preferences over wages and non-wage amenities. Let $W_j(X)$ be the wage that a firm j offers to workers with skills X, and let $a_j(X)$ be a firm's skill-specific amenity. For any worker i with skills X_i, the indirect utility from working at firm j is:

$$u_{ij}\left(W_j(X_i), a_j(X_i)\right) = \beta_i \log W_j(X_i) + a_j(X_i) + \epsilon_{ij}. \tag{1}$$

In this utility function, ϵ_{ij} represents the worker's idiosyncratic taste for firm j, reflecting non-pecuniary considerations such as distance from work and personal preferences over work environment. I assume that $\{\epsilon_{ij}\}_{i,j}$ are independent and identically distributed (i.i.d.) random variables, each following a continuous distribution function F_ϵ with positive density over $\mathbb{R}$.

The coefficient β_i is the worker's marginal utility of log earnings. This parameter governs the marginal rate of substitution between wages and amenities, thus capturing the trade-off between higher pay and better working conditions. A worker with a lower β_i is more willing to sacrifice an increase in earnings in exchange for improved non-wage job characteristics.[9] In Appendix A.8, I provide an explicit micro-foundation for the utility specification (1) where I re-interpret β_i as the worker's marginal rate of substitution between consumption and leisure.

In my analysis, I allow the coefficient β_i to vary freely among workers in the economy. Specifically, I treat β_i as a random variable that is distributed according to a density f_β. I do not place any parametric restrictions on this density, such as assuming a normal distribution for β_i. One important aspect of this framework is that I allow β_i to vary conditionally within each skill group X_i. As a result, equally skilled workers could exhibit different marginal rates of substitution with respect to wages and amenities. This level of variation is economically meaningful as it suggests that the trade-offs workers make when choosing jobs would reflect a variety of considerations beyond productivity. For example, β_i can vary based on contextual factors such as age, gender, household characteristics, and wealth. It may also be correlated among workers in the same household, neighborhood, or peer group. In my framework, I also

[8] This assumption reflects information asymmetries between workers and firms. It may also account for laws against pay discrimination, which are common in many labor markets. See Bhaskar et al. (2002) for discussion.

[9] Wiswall & Zafar (2018) and Maestas et al. (2023) examine how variation in the wage coefficient within a worker's utility function directly reflects differences in the willingness to trade earnings for non-wage amenities.

allow β_i to be correlated with workers' skills X_i. I assume that $\mathrm{P}(\beta_i > 0 | X_i = X) = 1$ for all X.

II.C. Firm Production

Let $D_j(\chi, \varphi)$ denote the mass of workers with skills $X = (\chi, \varphi)$ demanded by a firm j. For each skill type $\chi \in \mathcal{X}$, the total efficiency units of labor at the firm are defined such that:

$$L_j^{\text{eff}}(\chi) = \int \varphi D_j(\chi, \varphi) d\varphi. \tag{2}$$

Define Y_j to be the firm's value added from hiring labor. The firm's production function is a constant elasticity of substitution (CES) composite of the effective labor for different skill types. This formulation permits imperfect substitutability between skill types and allows for decreasing returns to scale, so that output need not increase in equal proportion to labor inputs.

$$Y_j = T_j \left(\sum_{\chi \in \mathcal{X}} \theta_{j\chi} \left(L_j^{\text{eff}}(\chi) \right)^{\rho_j} \right)^{\frac{1-\alpha_j}{\rho_j}}. \tag{3}$$

In this production function, $T_j \in \mathbb{R}_{++}$ is the firm's total factor productivity (TFP), $\theta_{j\chi} \in \mathbb{R}_{++}$ is the firm-specific efficiency of skill type χ, $(1-\rho_j)^{-1} \in [0, \infty)$ is the elasticity of substitution between skill types at the firm, and $1-\alpha_j \in (0, 1]$ is the firm's returns to scale. For simplicity, my specification abstracts from capital and intermediate inputs, as well as the structure of the product market. Nevertheless, as I show in Appendix A.9, this production function could be rationalized under either a perfectly competitive or an imperfectly competitive product market.

By allowing the productivity and technology parameters $(T_j, \{\theta_{j\chi}\}_{\chi \in \mathcal{X}}, \rho_j, \alpha_j)$ to vary among firms, this framework is able to accommodate a wide range of employer characteristics across different industries. Specifically, firms can be differentially productive, as captured by variation in T_j. In addition, as $\{\theta_{j\chi}\}_{\chi \in \mathcal{X}}$ is heterogeneous, the relative efficiencies of various skill types can differ among firms. For example, certain employers may place more emphasis on a college degree than others. Meanwhile, the amount of substitutability between these skill types can also vary, as reflected by differences in ρ_j. For example, employers may have specific job requirements that affect the ability to interchange high- and low-educated workers. Finally, due to variation in α_j, firms can exhibit different returns to scale. Importantly, my analysis does not impose restrictions on the relationship between amenities, productivity, and technology. Rather, I allow these fundamentals to vary freely across firms in the economy.

To better understand the relationship between skill types and skill levels in the production function, it will be useful to define the standardized unit of labor at the firm to be N_j, where:

$$N_j = \left(\sum_{\chi \in \mathcal{X}} \theta_{j\chi} \left(\int \varphi D_j(\chi, \varphi) d\varphi \right)^{\rho_j} \right)^{1/\rho_j}. \tag{4}$$

For any combination of skills $X = (\chi, \varphi)$, the elasticity of N_j with respect to $D_j(\chi, \varphi)$ equals:

$$\frac{\partial \log N_j}{\partial \log D_j(\chi, \varphi)} = \underbrace{\frac{\theta_{j\chi} \left(L_j^{\text{eff}}(\chi) \right)^{\rho_j}}{\sum_{\chi' \in \mathcal{X}} \theta_{j\chi'} \left(L_j^{\text{eff}}(\chi') \right)^{\rho_j}}}_{\partial \log N_j / \partial \log L_j^{\text{eff}}(\chi)} \times \underbrace{\frac{\varphi D_j(\chi, \varphi)}{\int \varphi' D_j(\chi, \varphi') d\varphi'}}_{\partial \log L_j^{\text{eff}}(\chi) / \partial \log D_j(\chi, \varphi)}. \tag{5}$$

This framework nests both efficiency units of labor and CES specifications. Within any skill type χ, workers exhibit varying productivity, as measured by φ. In contrast, workers that have different skill types are viewed as imperfect substitutes, and their productivities vary in a way that is specific to each firm. Under this setup, two skill levels φ and φ' are only comparable among workers in the same skill type, but they are not directly comparable across skill types.

II.D. Characterization of an Equilibrium

In equilibrium, workers select employers to maximize utility, as defined in equation (1). Given a set of wage offers $\{W_k(X)\}_{k=1}^J$ and non-wage amenities $\{a_k(X)\}_{k=1}^J$, a worker with a preference parameter β and a skill profile X chooses to work for a firm j with probability:

$$\mathrm{P}\left(j(i) = j | \beta, X\right) = \int_{-\infty}^{\infty} \prod_{k \neq j} F_\epsilon \left(\beta \log \left(\frac{W_j(X)}{W_k(X)} \right) + a_j(X) - a_k(X) + \tilde{\epsilon} \right) f_\epsilon(\tilde{\epsilon}) d\tilde{\epsilon}. \tag{6}$$

Averaging across the realizations of β, the labor supply at firm j for workers with skills X is:

$$S_j(X) = \int \left(\int_{-\infty}^{\infty} \prod_{k \neq j} F_\epsilon \left(\beta \log \left(\frac{W_j(X)}{W_k(X)} \right) + a_j(X) - a_k(X) + \tilde{\epsilon} \right) f_\epsilon(\tilde{\epsilon}) d\tilde{\epsilon} \right) f_{\beta, X}(\beta, X) d\beta. \tag{7}$$

Given the firm-specific labor supply curves described in equation (7), each firm optimally assigns wages to workers in order to maximize profit. Since the firm cannot discriminate based on individual preferences (β_i, ϵ_{ij}), its wage-setting decisions are informed by the distributions of these parameters, specifically $\{F_{\beta|X}\}_X$ and F_ϵ. The firm's profit maximization problem is:

$$\max_{\{W_j(\chi,\varphi)\}_{\chi,\varphi}} T_j \left(\sum_{\chi \in \mathcal{X}} \theta_{j\chi} \left(\int \varphi D_j(\chi,\varphi) d\varphi \right)^{\rho_j} \right)^{\frac{1-\alpha_j}{\rho_j}} - \sum_{\chi \in \mathcal{X}} \left(\int W_j(\chi,\varphi) D_j(\chi,\varphi) d\varphi \right), \quad (8)$$

subject to the constraint that labor demand equals supply: $D_j(X) = S_j(X)$ for all $X = (\chi, \varphi)$.

I consider an equilibrium in which firms view themselves to be strategically small in the economy. Consequently, a firm does not internalize the effect of changing its own wages on the labor that is supplied to other firms.[10] It is important to note that this assumption does not mean that workers regard firms as infinitesimal. On the contrary, there is a finite number of firms, and every worker chooses his/her job by considering the full set of wage offers and amenities. However, my analysis assumes that each firm sees itself as a marginal player, thereby ruling out strategic interactions in wage-setting. An equilibrium is defined as follows:

Definition. Given any set of worker skill and preference distributions $(F_{\chi,\varphi}, \{F_{\beta|\chi,\varphi}\}_{\chi,\varphi}, F_\epsilon)$, firm amenities $\{a_j(\chi,\varphi)\}_{\chi,\varphi}$ and production parameters $(T_j, \{\theta_{j\chi}\}_\chi, \rho_j, \alpha_j)$, an equilibrium is defined by the job decisions $j(i)$, labor supply curves $S_j(\chi,\varphi)$, and wages $W_j(\chi,\varphi)$ where:

(i) Each worker chooses a firm that maximizes his/her utility, as specified by equation (1).

(ii) The labor supply curves $S_j(\chi,\varphi)$ are consistent with workers' optimal choices, as in (7).

(iii) Each firm posts wages $W_j(\chi,\varphi)$ for all $\chi \in \mathcal{X}$ and $\varphi \in \mathbb{R}_+$ to maximize profit as defined in equation (8), ensuring that labor demand $D_j(\chi,\varphi)$ is equal to labor supply $S_j(\chi,\varphi)$.

In Lemma 1 of Appendix B.2, I establish the existence of an equilibrium involving strictly positive wages and employment. Lemma 2 further shows that this equilibrium is unique with probability one. Additionally, I analyze properties such as the concavity of the profit function and the conditions required for dynamic stability of the equilibrium. These properties will be important for running counterfactual analyses because they provide me with a way to compute equilibrium outcomes based on the underlying structural parameters in the model.

II.E. Properties of the Labor Supply Curve to a Firm

A primary motivation for including random coefficients in the worker's utility function is that it allows the model to capture rich substitution patterns that reflect sorting on individual heterogeneity. This subsection examines how this heterogeneity expands the scope of worker

[10] This assumption is also made by Card et al. (2018), Kroft et al. (2021), Lamadon et al. (2022), and Azar et al. (2022). Alternatively, Berger et al. (2022), Deb et al. (2024), and Chan et al. (2024) study models of job differentiation where firms internalize market share effects, and Jarosch et al. (2024) studies a search model with large firms. Identification is very difficult with strategic interactions, as it relies on more exclusion restrictions.

behavior in the model and what it implies about the shape of the labor supply curve to a firm.

Substitution Patterns under Homogeneous Effects

The standard assumption in models of job differentiation is that β is homogeneous among workers. Under this assumption, employer wages and amenities do not interact with worker-level preference parameters in the utility function. This additive separability between firm and worker characteristics makes the model more tractable. Yet, it imposes a strong form of homogeneity on aggregate substitution patterns that can lead to unrealistic market outcomes. This critique has previously been raised in the context of product markets (e.g., Berry et al., 1995). However, as I will demonstrate, it also holds significant implications for labor markets.

To illustrate how the homogeneous effects assumption restricts economic behavior, consider a scenario where β is constant across workers. In this case, the only reason that different workers rank firms differently is that their taste shocks ϵ_{ij} vary. Whenever these taste shocks are i.i.d., a worker's second choice firm does not depend on where that worker is employed.

This aspect of the model matters when analyzing a change in the economic environment. Specifically, if wages were to shift in the market, then a firm's employment response would not depend on that firm's characteristics. One way to formalize this idea is to consider two arbitrary firms, j and j'. I then analyze how the labor supplied to a third set of firms, defined as $S_{\mathcal{J}^*}(X) = \sum_{k \in \mathcal{J}^*} S_k(X)$ where $j, j' \notin \mathcal{J}^*$, is affected by the wages at j and j', respectively. When β is homogeneous, these labor supply responses are identical. In particular, I find that:

$$\frac{\partial \log S_{\mathcal{J}^*}(X)}{\partial \log W_j(X)} = \frac{\partial \log S_{\mathcal{J}^*}(X)}{\partial \log W_{j'}(X)}, \quad \text{for any } \mathcal{J}^* \in \{1, \ldots, J\} \text{ where } j, j' \notin \mathcal{J}^*. \tag{9}$$

This restriction implies that the workers at firms j and j' would be just as likely to substitute toward a third firm k, regardless of which of the two firms is more similar to firm k. This contradicts the intuition that firms with similar attributes should exhibit larger substitution effects. To see how this restriction affects a firm's labor supply curve, consider the following property.

Property 1. If β is homogeneous, then firms encounter identical labor supply elasticities:

$$\frac{\partial \log S_j(X)}{\partial \log W_j(X)} = \frac{\partial \log S_{j'}(X)}{\partial \log W_{j'}(X)}, \quad \text{for all } j, j' \in \{1, \ldots, J\}.$$

Property 1 underscores the limitations of using utility specification (1) with homogeneous effects. For example, consider two firms that offer different wages and amenities. It would be natural to expect that these firms attract different types of workers who respond differently to changes in their earnings. So, the elasticities of labor supply would likely differ between these firms. Nevertheless, if β is homogeneous, then the model cannot account for this pattern. In particular, firms would always encounter the same labor supply elasticities, even when they exhibit vastly different technologies and work environments. Importantly, these implications do not depend on the specific distribution of the idiosyncratic taste shocks, e.g., logit or probit.

One way to relax these restrictions on substitution patterns is to assume that $\{\epsilon_{ij}\}_{i,j}$ are not i.i.d., but rather that ϵ_{ij} is correlated among firms within local labor markets, e.g., via a nested logit structure. This method is taken by Azar et al. (2022) and Lamadon et al. (2022), among others. It implies that workers are more willing to substitute toward firms within markets than across markets. However, within each market, substitution effects would still be independent of firm characteristics, leading to constant labor supply elasticities across firms.[11]

A second extension could be to allow β to vary based on observable worker characteristics such as education or experience.[12] However, a concern with this strategy is that firms are also likely to observe these characteristics and would thus factor them into a worker's skills X. As long as β is constant conditional on skills, the implications of Property 1 remain unaffected. In particular, for any skill group X, the labor supply elasticities would be uniform across firms.

Substitution Patterns under Heterogeneous Effects

If β is heterogeneous, then the labor supply curve to a firm reflects differences in workers' preferences over wages and amenities. For example, firms offering higher wages and fewer amenities are more likely to attract workers who place a greater value on wages, as indicated by larger values of β. Conversely, firms that provide lower wages but better amenities will tend to attract workers who prioritize amenities, as indicated by smaller values of β. Thus, firms that share similar characteristics are more likely to attract workers with similar preferences over wages and amenities, which would lead these firms to exhibit larger substitution effects.

This variation in worker composition affects the elasticities of labor supply to a firm. To see how, note that any firm-specific elasticity can be represented as the average of β-specific elasticities for workers employed at the firm. The next property formalizes this relationship.

Property 2. The elasticity of labor supply faced by a firm j has the following representation:

$$\frac{\partial \log S_j(X)}{\partial \log W_j(X)} = \mathrm{E}\left(\frac{\partial \log \mathrm{P}(j(i) = j|\beta_i, X_i)}{\partial \log W_j(X_i)} \,\middle|\, X_i = X, j(i) = j \right).$$

As this property indicates, the elasticity of labor supply to a firm is shaped by the composition

[11] In principle, nests can be defined more broadly so they do not coincide with markets. Yet, these definitions still rely on factors observable to the analyst, which is often undesirable. See Nevo (2000) for more discussion.

[12] Examples of this approach include Chan et al. (2024) and Deb et al. (2024). The former uses a multinomial logit discrete choice model and the latter uses a representative agent CES model. As shown in Verboven (1996), Proposition 2, both modeling approaches lead to demand functions that are isomorphic at the market level.

of the firm's workforce, which is determined by the wages and amenities that the firm provides. Therefore, when β is heterogeneous, the labor supply elasticities are directly linked to firm characteristics in a way that would not be captured under a model with homogeneous effects.

Another implication of heterogeneous effects is that, even if firms are strategically small within the economy, the labor supply elasticities are endogenous to wage changes at a firm. When a firm posts a higher wage, it attracts workers who are more responsive to wage changes, i.e., those with larger β's. As a result, the elasticity of labor supply to the firm would increase. To formalize this dynamic, I analyze the second derivative of the labor supply curve to a firm.

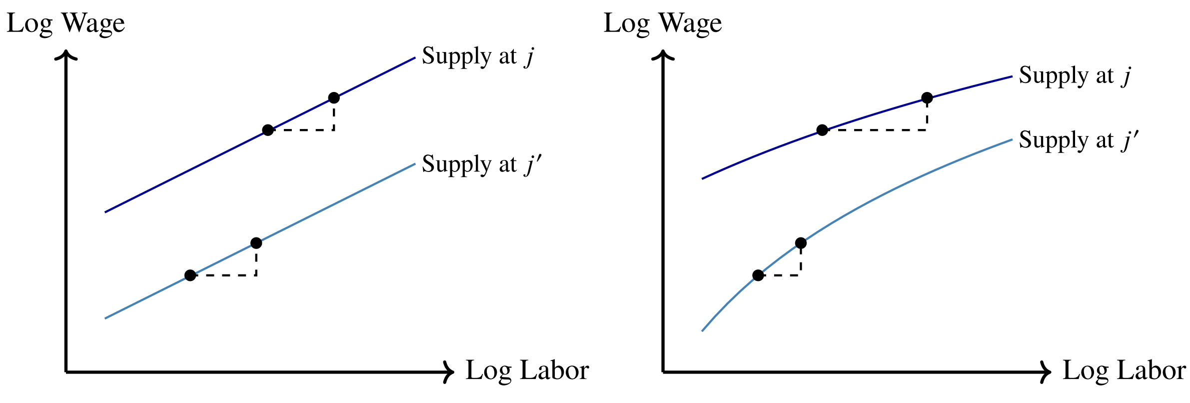
Property 3. The second derivative of log labor supply with respect to the log wage at a firm is:

$$\frac{\partial^2 \log S_j(X)}{\partial \left[\log W_j(X)\right]^2} = \mathrm{Var}\left(\frac{\partial \log \mathrm{P}(j(i) = j|\beta_i, X_i)}{\partial \log W_j(X_i)} \,\middle|\, X_i = X, j(i) = j\right).$$

This second derivative is strictly positive if and only if the variance $\mathrm{Var}(\beta_i|X_i = X)$ is nonzero.

Figure 1 illustrates how heterogeneous effects influence the labor supply curve to a firm. Consider two firms, j and j', where firm j has high wages but unfavorable amenities, while firm j' has low wages but favorable amenities. In the homogeneous effects case (depicted in Figure 1.a), both firms would face the same labor supply responses to internal wage changes, meaning that an X% wage change results in the same percent change in labor at either firm. Under heterogeneous effects (depicted in Figure 1.b), these two firms would face different labor supply responses due to differences in workforce composition. For the same X% wage change at each firm, the high-wage, low-amenity firm j would see a larger employment response, while the low-wage, high-amenity firm j' would see a smaller employment response.

Figure 1: Illustration of Firm-Specific Labor Supply Responses



(a) Homogeneous Effects (b) Heterogeneous Effects

Notes. This figure depicts the labor supply curves $\log S_j(X)$ and $\log S_{j'}(X)$ for two firms, j and j', respectively. Figure 1.a shows the case where $\mathrm{Var}(\beta_i \mid X_i = X) = 0$. Figure 1.b shows the case where $\mathrm{Var}(\beta_i \mid X_i = X) > 0$.

II.F. Wage Determination and Sorting in Equilibrium

In equilibrium, firms face upward-sloping labor supply curves, giving them the ability to set wages below the worker's marginal product of labor. Let $L_j(X)$ represent the amount of labor from skill group X employed at firm j in equilibrium, and define $\boldsymbol{\varepsilon}_j(X) = \frac{\partial \log L_j(X)}{\partial \log W_j(X)}$ to be the elasticity of labor supply faced by the firm. A worker's wage is determined as follows:

$$W_j(\chi, \varphi) = \underbrace{\frac{\boldsymbol{\varepsilon}_j(\chi, \varphi)}{1 + \boldsymbol{\varepsilon}_j(\chi, \varphi)}}_{\textit{Markdown}} \times \underbrace{\varphi T_j (1 - \alpha_j) \theta_{j\chi} \left(L_j^{\text{eff}}(\chi)\right)^{\rho_j - 1} \left(\sum_{\chi' \in X} \theta_{j\chi'} \left(L_j^{\text{eff}}(\chi')\right)^{\rho_j}\right)^{\frac{1-\alpha_j}{\rho_j} - 1}}_{\textit{Marginal Product of Labor}}. \quad (10)$$

There are three main factors that drive workers to systematically sort into different firms. One factor is that a worker's marginal product varies across firms. This variation arises from differences in firm-specific productivity parameters $(T_j, \{\theta_{j\chi}\}_\chi)$ and technology (ρ_j, α_j). In particular, productivity differences imply that the same worker will be more or less efficient depending on the firm, while technology differences affect how a worker's marginal product is influenced by the firm's existing labor inputs. Together, these features have the potential to generate large differences in worker productivity distributions both within and across firms.

A second factor influencing sorting is that firms have distinct non-wage amenities $a_j(X)$, which are allowed to potentially vary within a firm based on worker skills. These amenities imply that firms are vertically differentiated, meaning that some firms exhibit superior amenities compared to others. They also lead firms to be horizontally differentiated, since workers at the same firm may receive varying levels of benefit from the same amenity. Note that a firm's amenities may depend on its production parameters $(T_j, \{\theta_{j\chi}\}_\chi, \rho_j, \alpha_j)$, which means that these non-wage job characteristics could be directly linked to a worker's marginal product.

A third factor affecting sorting is the worker's wage markdown $\frac{\boldsymbol{\varepsilon}_j(X)}{1+\boldsymbol{\varepsilon}_j(X)}$. These markdowns differ across firms due to variation in workforce composition, reflecting an interaction between worker preferences and firm characteristics. For example, a low-productivity, high-amenity firm is likely to attract workers with smaller β's, who place less emphasis on wages relative to amenities. As such, the firm's workforce would be less responsive to wage changes. This aspect enables the firm to exert more wage-setting power, leading to larger markdowns. In my model, markdowns are specific to each skill group X, which reflects the possibility that skill groups have systematically different preferences over wages and amenities. Consequently, workers with different skills may experience different markdowns even within the same firm.

One notable implication of sorting in my model is that a worker's wage markdown tends to amplify the impact of firm productivity on earnings. To understand this point, consider that a firm's wage-setting power diminishes as its labor supply becomes more elastic. Also, as Property 3 demonstrates, the elasticity of labor supply to a firm increases with the firm's wage. Consequently, when a demand shock drives up wages at a firm, it also shifts the composition of the firm's workforce, making the labor supply more elastic. This larger elasticity reduces the firm's wage-setting power, resulting in even higher wages for workers. By this process, wage changes at a firm can become self-reinforcing.[13] Importantly, as I demonstrate in Appendix B, this feature of the model does not interfere with the uniqueness of the equilibrium.[14]

II.G. Rents, Pass-throughs, and Allocative Inefficiency

I end this section by examining how the structural parameters in the model correspond to different economic quantities, such as the rents earned by workers and firms, the pass-through of productivity shocks to earnings, and measures of welfare loss due to allocative inefficiency.

To simplify the derivation of these quantities, it will be useful to specify a functional form for the distribution of workers' taste shocks $\{\epsilon_{ij}\}_{i,j}$. To do so, I analyze a special case of the model where ϵ_{ij} has a Type 1 Extreme Value (Logit) distribution: $F_\epsilon(\epsilon) = \exp(-\exp(-\epsilon))$. This assumption does not place meaningful restrictions on the properties discussed in Section II.E, as these properties are primarily driven by the parameter β_i, not by ϵ_{ij}. Nevertheless, it allows me to derive clearer expressions for the firm's labor supply curve and its elasticities.

Under the logit model, a worker's choice probability, as defined in equation (6), becomes:

$$\mathrm{P}\left(j(i) = j | \beta, X\right) = \frac{\exp\left(\beta \log W_j(X) + a_j(X)\right)}{\sum_{k=1}^{J} \exp\left(\beta \log W_k(X) + a_k(X)\right)}. \tag{11}$$

Let $I(\beta, X) = \sum_{k=1}^{J} \exp\left(\beta \log\left(W_k(X)\right) + a_k(X)\right)$ denote the wage index for any worker with preference parameter β and skills X. This index varies across individuals, reflecting the fact

[13] Another way to visualize this effect is to analyze the wage elasticity of the markdown $\frac{\varepsilon_j(X)}{1+\varepsilon_j(X)}$, which equals:

$$\frac{\partial \log \frac{\varepsilon_j(X)}{1+\varepsilon_j(X)}}{\partial \log W_j(X)} = \frac{\frac{\partial^2 \log L_j(X)}{\partial [\log W_j(X)]^2}}{\varepsilon_j(X)[1+\varepsilon_j(X)]}, \quad \text{where} \quad \frac{\partial^2 \log L_j(X)}{\partial [\log W_j(X)]^2} \geq 0.$$

This elasticity is positive, indicating that fluctuations in wages at a firm are magnified through the markdown.

[14] In principle, the spillover effect of wages on markdowns can lead to multiple fixed points for the equilibrium wage equation (10). This multiplicity suggests that the firm's profit function might have two local optima: one characterized by lower wages, reduced employment, and larger markdowns; and the other marked by higher wages, increased employment, and smaller markdowns. This phenomenon is discussed by Manning (2010) in the context of spatial agglomeration. Fortunately, I find that multiplicity does not occur in my model as long as the conditional variance $\mathrm{Var}(\beta|X)$ is bounded in relation to the returns to scale $1 - \alpha_j$. I also prove that, even if multiple fixed points exist, there is, with probability one, a unique point where profit is globally maximized.

that workers have different reservation wages. The labor supply curve to a firm j is given by:

$$S_j(X) = \int \frac{1}{I(\beta, X)} \exp\big(\beta \log W_j(X) + a_j(X)\big) f_{\beta,X}(\beta, X) d\beta. \tag{12}$$

One key advantage of using this specification is that it leads to a clearer economic interpretation of the random coefficient β. In particular, under the logit error structure, β can be expressed as an individual-specific labor supply elasticity, defined as: $\frac{\partial \log \mathrm{P}(j(i)=j|\beta,X)}{\partial \log W_j(X)} = \beta$. This elasticity measures how a worker's probability of choosing a job at firm j is affected by variations in firm j's own wage. By Property 2, it follows that the labor supply elasticity faced by the firm represents a firm-wide average of the realizations of β. Specifically, I find:

$$\frac{\partial \log S_j(X)}{\partial \log W_j(X)} = \mathrm{E}\left(\beta_i \middle| X_i = X, j(i) = j\right). \tag{13}$$

Moreover, by Property 3, the curvature of a firm's labor supply curve—represented by its second derivative—can be defined in terms of the variance of β among workers at the firm:

$$\frac{\partial^2 \log S_j(X)}{\partial [\log W_j(X)]^2} = \mathrm{Var}\left(\beta_i \middle| X_i = X, j(i) = j\right). \tag{14}$$

In Appendix A.3, I also derive the third, fourth, and fifth derivatives of the labor supply curve to a firm under the logit model. I demonstrate that these derivatives are characterized by the higher-order moments of the distribution of β. In each of the derivations in the Appendix, I carefully highlight which parts of the analysis are reliant on utilizing the logit error structure.

Worker Rents

In this economy, firms cannot observe workers' individual preferences β_i and ε_{ij}, which means that they are unable to price-discriminate based on each worker's reservation wage. As a result, an equilibrium involves surpluses—or rents—for workers who are inframarginal. These surpluses represent the additional earnings that workers are willing to forgo to remain indifferent between their current job and their next-best alternative. Let R_i^w denote the rents that a worker i receives from his or her current job. A worker's rents are defined as follows:

$$u_{ij}\big(W_j(X_i) - R_i^w, a_j(X_i)\big) = \max_{j' \neq j(i)} u_{ij'}\big(W_{j'}(X_i), a_{j'}(X_i)\big). \tag{15}$$

To illustrate how worker rents systematically vary in the economy, I derive an expression for the expected rents $\bar{R}_{jX}^w = \mathrm{E}(R_i^w | j(i) = j, X_i = X)$ for workers with skills X employed at firm j.

$$\bar{R}_{jX}^w = W_j(X) \times \mathrm{E}\left(\frac{1}{1+\beta_i} \,\middle|\, j(i) = j, X_i = X\right). \tag{16}$$

In a model with homogeneous effects, where $\beta_i = \beta$ for all i, a worker's expected rents are a fixed proportion of his or her earnings, specifically: $\bar{R}^w_{jX}/W_j(X) = (1+\beta)^{-1}$. As a result, the model predicts that all workers, irrespective of their employer j or their skills X, would expect to receive the same share of their wages as rents. This restriction runs counter to the intuition that workers at different jobs would benefit in different ways from their wage contracts.

When β is heterogeneous, the share of wages that workers obtain as rents systematically varies across firms and skill groups. To see how this variation occur across firms, I examine how the ratio $\bar{R}^w_{jX}/W_j(X)$—which measures expected rents as a share of a worker's pay—is driven by firm characteristics. I find that this ratio is a decreasing function of a firm's wage:

$$\frac{\partial\left[\bar{R}^w_{jX}/W_j(X)\right]}{\partial \log W_j(X)} = \mathrm{Cov}\left(\beta_i, \frac{1}{1+\beta_i} \,\middle|\, j(i) = j, X_i = X\right) \leq 0. \tag{17}$$

This negative relationship reflects how workers sort based on variation in β. Within each skill group X, workers who choose jobs at higher-paying firms are more likely to have higher reservation wages. Thus, even while these workers earn more, they tend to receive a smaller share of their wages as rents. As I show empirically, neglecting this source of heterogeneity leads to an inaccurate assessment of where worker rents are concentrated within the economy.

Firm Rents

Firms also capture rents in equilibrium. These rents come in the form of excess profits that firms obtain by exploiting their wage-setting power. To quantify these rents, I analyze a counterfactual setting where a firm is a price-taker in the economy, facing a perfectly elastic labor supply curve. For any firm j, I construct the counterfactual profits $\Pi_j^{\text{price-taker}}$ by solving:

$$\max_{\{D_j^{\text{pt}}(\chi,\varphi)\}_{\chi,\varphi}} T_j\left(\sum_{\chi\in\mathcal{X}} \theta_{j\chi}\left(\int \varphi D_j^{\text{pt}}(\chi,\varphi)d\varphi\right)^{\rho_j}\right)^{\frac{1-\alpha_j}{\rho_j}} - \sum_{\chi\in\mathcal{X}}\left(\int W_j^{\text{pt}}(\chi,\varphi)D_j^{\text{pt}}(\chi,\varphi)d\varphi\right), \tag{18}$$

subject to the labor supply constraint: $D_j^{\text{pt}}(X) = S_j^{\text{pt}}(X)$ for all $X = (\chi,\varphi)$. In this counterfactual, the only difference in the firm's behavior is that it does not exert wage-setting power. Therefore, this counterfactual differs from one where all firms in the economy are price-takers.

I define a firm j's rents as the difference between actual profits and counterfactual profits at the firm, expressed as $R_j^e = \Pi_j - \Pi_j^{\text{price-taker}}$. In Appendix A.5, I show that these rents equal:

$$R_j^e = Y_j \times \left[\sum_{\chi\in\mathcal{X}}\left[\int\left(\frac{1+\alpha_j\,\boldsymbol{\varepsilon}_j(\chi,\varphi)}{1+\boldsymbol{\varepsilon}_j(\chi,\varphi)}\right)\frac{\partial \log N_j}{\partial \log L_j(\chi,\varphi)}d\varphi\right] - \alpha_j\left(\frac{Y_j^{\text{pt}}}{Y_j}\right)\right], \tag{19}$$

where Y_j^{pt} represents the firm's value added as a price-taker. In my model, Y_j^{pt} does not appear to have a closed-form expression. Nevertheless, I show how to solve for this counterfactual quantity from the structural parameters in the model using a gradient descent algorithm. This method relies on the existence of a unique solution to the price-taker's first-order condition and the concavity of the firm's profit function, both of which are established in Appendix B.2.

Under my framework, the relationship between a firm's productivity and its rents R_j^e is ambiguous. On the one hand, a more productive firm generates greater value added, thereby increasing profitability. On the other hand, increased productivity typically leads to higher wages, which in turn results in higher labor supply elasticities $\{\boldsymbol{\varepsilon}_j(X)\}_X$ faced by the firm. As these labor supply elasticities become larger, the firm's market power declines, making the labor inputs even more costly. In particular, as $\boldsymbol{\varepsilon}_j(X) \to \infty$ for all X, the firm's profits will converge to those of a price-taking firm, which causes the rents to approach zero: $R_j^e \to 0$.

A firm's rents are also influenced by its technology. Specifically, as the firm approaches constant returns to scale, i.e., $\alpha_j \to 0$, the profits it could earn as a price-taker tend to zero. So, any profits that the firm manages to capture will come in the form of rents: $\lim_{\alpha_j \to 0} \Pi_j = R_j^e$.

In my analysis, it is hard to empirically distinguish between a firm's ex-ante and ex-post rents. This task would require knowing more about how firms select amenities, which I do not observe in the data. Therefore, for identification, I assume that firms are unable to manipulate their amenities—at least during the time frame that the model is estimated. Importantly, this assumption does not restrict how amenities correlate with a firm's productivity or technology. It also does not rule out the possibility that firms initially choose amenities to maximize profits.

Pass-through of TFP Shocks to Wages

Next, I examine the impact of a hypothetical shock to firm productivity—specifically, the TFP parameter T_j—on workers' wages. This reduced form quantity is commonly used to learn about rent-sharing, as it indicates how changes in a firm's value added are passed on to workers in the form of pay adjustments.[15] I derive the pass-through of TFP shocks to wages as:

$$\frac{\partial \log W_j(\chi,\varphi)}{\partial \log T_j} = \Bigg[1 - \left(\frac{1}{\boldsymbol{\varepsilon}_j(\chi,\varphi)[1+\boldsymbol{\varepsilon}_j(\chi,\varphi)]}\right) \times \frac{\partial^2 \log L_j(\chi,\varphi)}{\partial[\log W_j(\chi,\varphi)]^2} + (1-\rho_j)\int \boldsymbol{\varepsilon}_j(\chi,\varphi')\frac{\partial \log L_j^{\text{eff}}(\chi)}{\partial \log L_j(\chi,\varphi')}d\varphi' - (1-\alpha_j-\rho_j)\sum_{\chi'\in\mathcal{X}}\left(\int \boldsymbol{\varepsilon}_j(\chi',\varphi')\frac{\partial \log N_j}{\partial \log L_j(\chi',\varphi')}d\varphi'\right)\Bigg]^{-1}. \quad (20)$$

[15]For example, Card et al. (2018) explicitly interprets the pass-through $\frac{\partial \log W_j(X)}{\partial \log T_j}$ as a rent-sharing elasticity.

To interpret the pass-through, first consider a case where β is homogenous across workers. In this setting, the firm's labor supply curve is isoelastic, which implies that $\boldsymbol{\varepsilon}_j(\chi, \varphi) = \boldsymbol{\varepsilon}$ for a scalar $\boldsymbol{\varepsilon} \in \mathbb{R}_+$ and $\frac{\partial^2 \log L_j(\chi,\varphi)}{\partial[\log W_j(\chi,\varphi)]^2} = 0$. Under these restrictions, the pass-through reduces to:

$$\frac{\partial \log W_j(\chi, \varphi)}{\partial \log T_j} = \frac{1}{1 + \alpha_j \boldsymbol{\varepsilon}}. \tag{21}$$

This pass-through varies across firms in the economy due differences in the returns to scale parameters $\{\alpha_j\}_j$, with firms that have higher returns to scale exhibiting larger pass-throughs.

When β is heterogeneous among workers, the pass-through changes in two key aspects. First, since the labor supply elasticity $\boldsymbol{\varepsilon}_j(\chi, \varphi)$ may differ by skill type χ, the pass-through will depend on the elasticity of substitution between skill types at a firm, specifically $(1 - \rho_j)^{-1}$. This feature introduces an additional source of variation across firms. Second, since β varies among equally-skilled workers, the wage markdown at a firm will be endogenous to changes in the firm's productivity. Specifically, a positive TFP shock that increases wages will alter the firm's workforce composition in a way that reduces its wage-setting power, resulting in even higher wages at the firm. This effect is captured by the positive second derivative of the labor supply curve, given by $\frac{\partial^2 \log L_j(X)}{\partial[\log W_j(X)]^2}$. Note that, in principle, the pass-through $\frac{\partial \log W_j(X)}{\partial \log T_j}$ can be greater than one, although my empirical results show that it is less than one in practice.

Welfare and Allocative Inefficiency

I study a measure of social welfare that incorporates firm profits. Specifically, I assume that profits are redistributed to workers in proportion to their wages. This approach ensures that the redistribution does not distort workers' employment decisions. Welfare is defined as:

$$\mathcal{W} = \mathrm{E}\left(\max_j \left\{u_{ij}\big(W_j(X_i), a_j(X_i)\big)\right\}\right) + \log\left(\sum_{j=1}^{J} \Pi_j\right). \tag{22}$$

To evaluate the welfare cost of imperfect competition, I compare welfare in the monopsonistic economy to that which would be obtained by a social planner who optimally allocates labor to firms by solving: $\mathcal{W}^* = \max_{\{j(i)\}_i} \mathcal{W}$. I assume that the social planner has complete knowledge of worker and firm fundamentals but that firms continue to set wages based on their limited information of (β_i, ϵ_{ij}). As I prove in Appendix A.7, the planner's optimal allocation is equivalent to what would be achieved in a competitive (Walrasian) economy, where firms act as price takers facing perfectly elastic labor supply curves. Moreover, I show that it is possible to solve for welfare under this optimal allocation from the model's structural parameters.

In an economy with constant markdowns, there would be no misallocation of labor due to imperfect competition, since workers would sort into the same firms that they would have chosen in the absence of monopsony power. However, if markdowns vary across firms, then there is a potential for allocative inefficiency. This inefficiency occurs, for example, when a worker who is most productive at one firm j earns a higher wage at another firm k because firm j has a larger markdown than firm k. These discrepancies can alter workers' rankings of firms and distort their labor supply decisions. In this way, heterogeneous markdowns introduce a welfare cost of imperfect competition that is not present in the case of constant markdowns.[16]

III. Data and Empirical Context

I now describe the data sources and key variables that are used in my empirical analysis, as well as the criteria for sample selection. Additional details are provided in the Appendix.

III.A. Data Sources

My analysis uses a matched employer-employee panel dataset from Norway, containing information about workers and firms over the period 1995-2018. The data is constructed by linking annual business tax records to a national employment register from the Norwegian Labour and Welfare Administration. It encompasses all workers who are not self-employed and it covers the universe of firms that are required to file a tax return. Further information on worker demographics and financial assets come from individual tax filings and social security registers. All administrative data is maintained by Statistics Norway and is subject to comprehensive quality control measures. This data is considered to be highly reliable.[17]

At the worker-level, the data contains information on annual earnings, scheduled work hours, and the duration of a job spell. Earnings account for a worker's salary, bonuses, and other payments offered by an employer such as overtime, vacation, and severance pay.[18] All monetary quantities are adjusted for inflation and are expressed in 2014 US dollars. The data also includes details about a worker's education, experience, and specialization (e.g., college major and vocational training). These variables are used to define a worker's skill type χ. In the baseline specification, I define χ as a composite of education and experience, where a worker's education level is divided into three categories (no high school degree, high school degree but no college degree, and college degree) and experience is divided into two categories (<10 years in the workforce and ≥ 10 years in the workforce). Alternate definitions of χ are considered in the Appendix. Lastly, the data contains rich demographic information on workers and their families. These data will be used for making cross-sectional comparisons.

In the data, workers are linked to establishments, which represent sub-units (or branches) of a firm. To address this dimension, I measure employment at the establishment-level, and I

[16]This conclusion assumes that there is no nonemployment margin. In a model involving nonemployment, misallocation can exist even with constant markdowns, as the trade-off between labor and leisure is distorted.

[17]See Fagereng et al. (2021) for discussion. Bruun et al. (2021) also give an independent quality evaluation.

[18]In-kind benefits are also included in earnings; however, parental leave and sickness benefits are excluded.

assume that a firm's production function aggregates labor in a way that allows for imperfect substitutability between establishments.[19] This approach enables greater flexibility in sorting and wage determination. For example, if an oil and gas firm operates a processing plant in rural Norway and has a corporate office in Oslo, then these two establishments are treated as distinct entities—both by workers and by the managing firm. Establishments are assigned to local labor markets based on industry and geographic location. The Central Register of Establishments and Enterprises identifies 10 broad industries and 46 commuting zones, as defined in Bhuller (2009). A market is constructed as the intersection of these two identifiers.

Accounting data is collected at the firm-level based on corporate income statements and balance sheets. A firm's value-added is defined as the difference between revenue and non-labor expenses, which include operating costs, spending on intermediate inputs, and capital depreciation. Profit is calculated to be revenue minus total costs, where total costs account for the wages paid to workers. All these quantities are measured prior to interest and taxes.

III.B. Sample Construction

As my framework studies labor supply on the extensive margin, I restrict the estimation sample to full-time workers, defined as those who work at least 30 hours per week at a given establishment. Throughout my analysis, labor will be measured as full-time employment. To calculate earnings, I further restrict the sample to individuals aged 25 to 55 who work at the same establishment for an entire year and who do not have a higher-paying job elsewhere. Following Song et al. (2018), I also require that a worker's earnings are above a minimum threshold. In the baseline sample, this threshold is set to \$15,000 per year. In the Appendix, I show that my empirical results are robust to changes in this minimum earnings threshold.

Part of my analysis relies on utilizing a balanced panel of firms. To achieve this objective, I limit the sample to firms that remain operational for nine consecutive years. During these years, each firm must consistently report positive revenue and maintain at least five full-time employees. In addition, each firm must continue operating the same set of establishments, and each establishment must stay in the same industry and commuting zone for the full nine years.

Table 1 provides descriptive statistics for the main estimation sample. More information about the distribution of wages, labor, value added, and profits can be found in the Appendix.

[19] Specifically, if a firm j operates multiple establishments $k \in \mathcal{K}_j$, then I adapt the production function so that $Y_j = T_j [\sum_{k \in \mathcal{K}_j} \sum_{\chi \in \mathcal{X}} \theta_{jk\chi} (L^{\text{eff}}_{jk}(\chi))^{\rho_j}]^{(1-\alpha_j)/\rho_j}$, where $L^{\text{eff}}_{jk}(\chi) = \int \varphi D_{jk}(\chi, \varphi) d\varphi$ is the total efficiency units of labor demanded by firm j for establishment k. This extension does not meaningfully change the theoretical results or the identification analysis. Therefore, to ease notation, the ongoing discussion will primarily focus on a setting where each firm operates only one establishment, i.e., $|\mathcal{K}_j| = 1$. As an additional robustness check in the Appendix, I consider the sensitivity of my estimates to restricting the sample to single-establishment firms.

Table 1: Summary Statistics—Main Estimation Sample

	Unique Identifiers	Unique Observations
Panel A. Population Counts		
Number of Full-time Workers	1,325,863	8,303,075
Number of Establishments	31,252	475,614
Number of Firms	23,066	274,687
Number of Local Markets	457	10,563
	Share of Workforce	**Mean Log Earnings**
Panel B. Worker Skill Types		
Lower Secondary, <10 Years Exp.	6.87%	10.91
Upper Secondary, <10 Years Exp.	18.98%	11.11
College Graduate, <10 Years Exp.	14.74%	11.33
Lower Secondary, ≥10 Years Exp.	21.30%	11.09
Upper Secondary, ≥10 Years Exp.	23.67%	11.30
College Graduate, ≥10 Years Exp.	14.43%	11.59
	Mean	**Std. Dev.**
Panel C. Accounting Variables		
Log Revenues	18.05	2.20
Log Total Costs	18.01	2.19
Log Value Added	16.86	2.18
Log Profit	14.97	2.54

Notes. This table presents descriptive statistics for the baseline estimation sample. All monetary quantities are adjusted for inflation and are expressed in log 2014 US dollars.

IV. Identification of Economic Quantities

In this section, I demonstrate how the model can be taken to data. Specifically, I show how to learn about key economic quantities, such as markdowns, rents, and allocative inefficiency, from the matched employer-employee dataset described in Section III. While some of these quantities can be recovered without full knowledge of the model's structural parameters, others—particularly those defined by counterfactual outcomes—do rely on the entire set of parameters. For this reason, my objective is to achieve full point identification of the model.

My identification analysis follows two main steps. First, I demonstrate how to recover the firm-specific labor supply elasticities $\boldsymbol{\varepsilon}_j(\chi, \varphi)$ through the use of panel data by leveraging an economic shock that affects the productivity of multiple firms in the economy. This strategy relies on instrumental variables, as well as assumptions about the unobserved skill levels φ that enable me to distinguish between the labor from different skill groups at a given firm. Second, I prove the identification of firms' production parameters $(T_j, \{\theta_{j\chi}\}_\chi, \rho_j, \alpha_j)$ and amenities $\{a_j(\chi, \varphi)\}_{\chi,\varphi}$, as well as the conditional densities $\{f_{\beta|\chi,\varphi}\}_{\chi,\varphi}$ corresponding to the distribution of workers' preferences over wages and amenities. As I demonstrate, the ability to determine these structural quantities relies on credibly recovering the elasticities $\boldsymbol{\varepsilon}_j(\chi, \varphi)$.

IV.A. *Identification of Firm-Specific Labor Supply Elasticities*

IV.A.1. *Ideal Experiment*

I begin by laying out the comparative statics that enable the identification of labor supply elasticities. Consider a hypothetical productivity shock that shifts a firm j's TFP in logs from $\log(T_j)$ to $\log(T_j)+\delta_j$. Assume that all other parameters in the economy remain unchanged. For each skill group X, let $\big(W_j(X), L_j(X)\big)$ and $\big(W'_j(X), L'_j(X)\big)$ represent the equilibrium wage and employment levels at the firm before and after the TFP shock, respectively. As δ_j approaches zero, these equilibrium quantities can be used to recover the elasticity $\boldsymbol{\varepsilon}_j(X)$. In particular, as I demonstrate in the Appendix, the equilibrium outcomes satisfy the condition:

$$\lim_{\delta_j \to 0} \Big[\big(\log L'_j(X) - \log L_j(X) \big) - \boldsymbol{\varepsilon}_j(X) \times \big(\log W'_j(X) - \log W_j(X) \big) \Big] = 0. \tag{23}$$

This condition follows from two key properties of the model. First, in the equilibrium characterization, firms view each other as strategically small. Therefore, they do not react to wage changes that are confined to a single firm in the economy. As a result, the labor supply $L_j(X)$ to a firm j would only be affected by an exogenous shock to T_j through changes in firm j's own wage $W_j(X)$. The second property of the model is that the equilibrium is locally stable, which means that iterating on the equilibrium conditions (10) converges to an equilibrium for all nearby initial values. This property is important for comparative statics as it implies that a small productivity shock does not lead to large jumps in wages or employment.

Based on equation (23), an ideal experiment to recover the elasticities $\{\boldsymbol{\varepsilon}_j(X)\}_{j,X}$ would involve inducing independent and infinitesimal shocks to each firm's TFP and then measuring how wages and labor respond. However, such an experiment is not practically feasible. In particular, I cannot independently manipulate each firm's productivity. Moreover, even if it were possible to do so, I cannot directly observe a worker's skill level φ. Thus, I would be unable to fully differentiate between labor inputs at a firm. In the remainder of this subsection, I explain how to overcome these two challenges using panel data and instrumental variables.

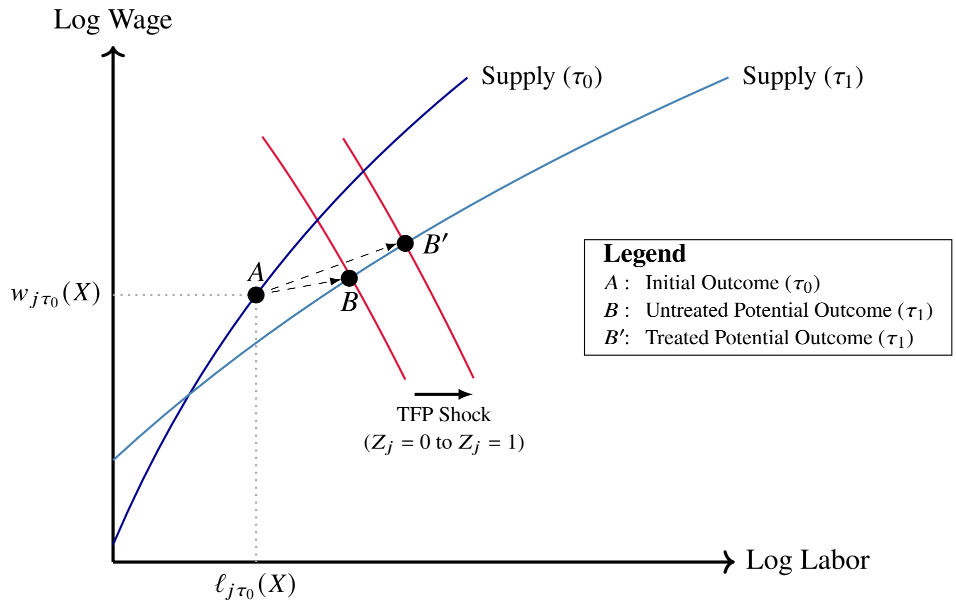
IV.A.2. *Econometric Model and Assumptions*

I now describe the empirical framework and assumptions that form the basis for my identification results. Throughout this discussion, I will utilize the logit error structure, where I assume $F_\epsilon(\epsilon) = \exp(-\exp(-\epsilon))$. Consequently, the firm-specific labor supply curves will be defined by equation (12). To ease notation, I use lowercase letters to denote quantities that are in log form, specifically defining $w_j(X) = \log W_j(X)$, $\ell_j(X) = \log L_j(X)$, and $y_j = \log Y_j$.

Description of Empirical Framework for a Generic Instrument

Consider the evolution of equilibrium outcomes over two time periods: τ_0 and τ_1. After the initial period τ_0, an event occurs whereby a subset of firms in the economy experiences a TFP shock, while the productivity of other firms remains unchanged. Following this event, wages and labor adjust to a new equilibrium by the second period τ_1. Let $\{w_{j\tau_0}(X), \ell_{j\tau_0}(X)\}_X$ denote the log wages and labor for a firm j in period τ_0, and let $\{w_{j\tau_1}(X), \ell_{j\tau_1}(X)\}_X$ be the corresponding values in τ_1. Additionally, let Z_j be a binary variable indicating whether firm j experiences a TFP shock. Firms are classified as *treated* if $Z_j = 1$ and *untreated* if $Z_j = 0$.

Figure 2: Illustration of Potential Outcomes for a Firm



Notes. This figure depicts how a firm's potential outcomes evolve in response to an event that shifts the TFP of firms in the economy. At τ_0, worker wage indices, skills and preferences are characterized by $\{I_{\tau_0}(\beta, X)\}_{\beta,X}$ and $\{F_{\beta,X|\tau_0}\}_X$, respectively. At τ_1, they are characterized by $\{I_{\tau_1}(\beta, X)\}_{\beta,X}$ and $\{F_{\beta,X|\tau_1}\}_X$, respectively. Aside from the TFP parameter T_j, which shifts if firm j is treated, the firm's characteristics $(\{\theta_{j\chi}\}_\chi, \rho_j, \alpha_j, \{a_j(X)\}_X)$ are constant over time.

Figure 2 illustrates the evolution of wages and labor for a given firm j, delineating which aspects of the model are constant and which vary over time. Crucially, between periods τ_0 and τ_1, wages and labor adjust for all firms, including those that are untreated. To understand why, note that the TFP shocks are not limited to a single firm, but rather they occur at multiple firms in the economy. Consequently, wage changes among the treated firms would shift the wage indices $I(\beta, X)$ of workers, which in turn reshapes the labor supply curve to every firm. Therefore, the TFP shocks generate spillover effects that impact all employers in the economy.

Contributing to these dynamics, I allow the skills and preferences of workers to change over time in response to the productivity shocks affecting firms. This variation implies that the joint distribution of workers' skills $F_{\chi,\varphi}$, along with the conditional distributions $\{F_{\beta|\chi,\varphi}\}_{\chi,\varphi}$ that govern workers' preferences over wages and amenities, can evolve freely between periods τ_0 and τ_1 due to shifting economic factors arising from the previously described event. While I also allow workers' idiosyncratic tastes to evolve, I assume that the cross-sectional distribution of these tastes retains a logit structure in each period, i.e., $F_\epsilon(\epsilon_{ij\tau}) = \exp(-\exp(-\epsilon_{ij\tau}))$. To achieve this, I assume that $\epsilon_{ij\tau}$ is a Markov process with independent innovations across i.

Given the evolution of workers' wage indices, skills, and preferences over time, firms face different labor supply curves in periods τ_0 and τ_1. Thus, in order to measure the elasticity of labor supply to a firm j, it is necessary to compare the firm's potential outcomes within the same time period. In period τ_0, the treated and untreated potential outcomes are equal, since no TFP shocks have yet occurred. However, in period τ_1, the potential outcomes differ, as illustrated by points B and B' in Figure 2. This difference reflects the comparative statics from the ideal experiment, indicating how firm j's wages and labor would respond to a TFP shock that is confined only to that firm. If it were possible to compare both potential outcomes for a firm in period τ_1, then one could determine the firm's labor supply elasticity in that period.

Identification Assumptions

In my identification analysis, I will impose two key assumptions. First, I restrict how the exposure to a TFP shock, represented as Z_j, depends on the inherent characteristics of a firm.

Assumption I (*Instrument Independence*)**.** $Z_j \perp \big(\rho_j, \alpha_j, \{\theta_{j\chi}\}_\chi, \{a_j(X) - a_j(X')\}_{X,X'}\big)$.

Assumption I states that the firm's treatment status Z_j is statistically independent of its technology (ρ_j, α_j), the relative efficiencies $\{\theta_{j\chi}\}_\chi$ of each skill type, and the differences in amenities $\{a_j(X) - a_j(X')\}_{X,X'}$ between skill groups at the firm. Importantly, this assumption does not require Z_j to be fully exogenous. In particular, I allow it to depend on a firm's productivity through the baseline TFP T_j. Additionally, while Z_j is independent of amenity differences within a firm, it may still be affected by a firm's average amenity $\mathrm{E}(a_j(X)|j)$. Consequently, treated and untreated firms can exhibit systematically different wages and labor.

Next, I make an assumption about the relationship between a worker's skill level φ and the other structural parameters in the model. This assumption consists of three main restrictions.

Assumption II (*Restrictions on Unobserved Skill Levels*)**.**

II.1. $\varphi \perp \beta | \chi, \tau$.
II.2. $a_j(\chi, \varphi) = a_{j\chi} + a_{\chi\varphi}$.
II.3. $\mathrm{E}(\log \varphi | \chi, \tau) = \mathrm{E}(\log \varphi | \chi)$.

Assumption II.1 asserts that, given a worker's skill type χ and time period τ, preferences over wages and amenities, as captured by the parameter β, do not depend on the skill level φ. This assumption does not imply that β is fully independent of worker skills, as I still allow β

to depend on the skill type χ. Additionally, it does not rule out the possibility that β varies along unobserved skill dimensions; however, it does preclude systematic variation based on φ.

Assumption II.2 states that the non-wage amenity $a_j(\chi, \varphi)$ can be expressed as the sum of two components: one that is specific to the firm and the skill type, $a_{j\chi}$, and another that is specific to the skill level, $a_{\chi\varphi}$. This assumption ensures that amenities do not vary across firms with respect to unobserved aspects of worker skills. To better interpret the two amenity components $a_{j\chi}$ and $a_{\chi\varphi}$, I impose the normalization $\mathrm{E}(a_{\chi\varphi}|\chi) = 0$ without loss of generality. This normalization allows me to define $a_{j\chi}$ as the average amenity at firm j for skill type χ.

Assumption II.3 requires that the average log skill level $\mathrm{E}(\log\varphi|\chi, \tau)$ among workers in a given skill type χ is invariant over time τ. Note that this assumption does not restrict how the variance of these skill levels, or other higher-order moments, evolve. Moreover, it does not restrict the evolution of the distribution $F_{\chi|\tau}$, which governs workers' observable skill types. In my subsequent analysis, it will be useful to normalize φ by setting $\mathrm{E}(\log\varphi|\chi) = 0$ for each $\chi \in \mathcal{X}$. This normalization is without loss of generality, as it is always possible to redefine the relative efficiencies $\{\theta_{j\chi}\}_{j,\chi}$ of each skill type to include the values of $\{\mathrm{E}(\log\varphi|\chi)\}_{\chi}$.[20]

IV.A.3. Identification Results

I now provide results on the identification of firm-specific labor supply elasticities. First, I establish the identification of unobserved skill levels. Next, I outline the IV design used to recover the elasticities. Finally, I discuss the sources of instruments employed in my analysis.

Identification of Worker Skills

To draw inferences about the labor supply elasticities, I will need to distinguish between workers with different skills. These skills are only partially observed in the data, however, as I do not directly measure a worker's skill level φ. To overcome this challenge, I rely on the restrictions implied by Assumption II. I demonstrate that, under this assumption, the skill levels φ are identified up to scale from observed differences in wages among workers at a given firm.

To establish identification, I first examine the implications of Assumption II for workers' labor supply decisions and the determination of wages at a firm. The key implication of this assumption is that the labor supply elasticities faced by a firm j do not vary along unobserved skill dimensions, i.e., $\boldsymbol{\varepsilon}_{j\tau}(\chi, \varphi) = \boldsymbol{\varepsilon}_{j\tau}(\chi)$. Thus, the wage markdowns at a firm j differ by workers' skill types χ, but not by their skill levels φ. This property implies the following result.

Proposition 1. Under Assumption II, the wage can be written as: $W_{j\tau}(\chi_i, \varphi_i) = \varphi_{i\tau} \times \psi_{j\chi\tau}$, where the term $\psi_{j\chi\tau}$ is uniform across workers in a given firm j, skill type χ, and period τ.

[20] Specifically, if $\mathrm{E}(\log\varphi|\chi) \neq 0$ for some skill type χ, then I can redefine $\log\theta_{j\chi}$ as $\log\theta_{j\chi} + \rho_j\,\mathrm{E}(\log\varphi|\chi)$.

This result provides a way to compute workers' skill levels from data about workers' earnings. In particular, I can recover φ from cross-sectional wage differences in each firm and skill type:

$$\log \varphi_{i\tau} = \log W_{j\tau}(\chi_i, \varphi_i) - \mathrm{E}\left[\log W_{j\tau}(\chi_i, \varphi_i) \mid j, \chi, \tau\right]. \tag{24}$$

Going forward, it will be useful to measure wages in terms of efficiency units. I define the effective wage for workers with skill type χ at firm j in period τ as $W^{\text{eff}}_{j\tau}(\chi) = W_{j\tau}(\chi, \varphi)/\varphi$.[21] Using this notation, I can re-interpret the elasticity of labor supply to a firm as a function of the firm's effective wage. Specifically, I demonstrate that $\boldsymbol{\varepsilon}_{j\tau}(\chi)$ takes the following form:

$$\boldsymbol{\varepsilon}_{j\tau}(\chi) = \int \beta \times \frac{\left[\exp\left(\beta \log W^{\text{eff}}_{j\tau}(\chi)\right) / I^{\text{eff}}_{\tau}(\beta, \chi)\right] f_{\beta|\chi,\tau}(\beta|\chi, \tau)}{\int \left[\exp\left(\beta' \log W^{\text{eff}}_{j\tau}(\chi)\right) / I^{\text{eff}}_{\tau}(\beta', \chi)\right] f_{\beta|\chi,\tau}(\beta'|\chi, \tau) d\beta'} d\beta. \tag{25}$$

In this expression, $I^{\text{eff}}_{\tau}(\beta, \chi) = \sum_{k=1}^{J} \exp(\beta \log W^{\text{eff}}_{k\tau}(\chi) + a_{k\chi})$ can be interpreted as the effective wage index for workers with preference parameter β and skill type χ during time period τ.

Difference-in-Differences Estimand

I now present the main identification result of this subsection, demonstrating how to recover the elasticity of labor supply to a firm using a difference-in-differences (DiD) research design. To motivate my approach, it is first important to recognize that I would not be able to determine the elasticities through either a cross-sectional comparison of treated and untreated firms or a comparison of outcomes within a firm over time. A cross-sectional comparison is not suitable because treated and untreated firms face different labor supply curves due to systematic variation in amenities. Moreover, a within-firm panel regression fails because of the spillover effects of TFP shocks on workers' wage indices and preferences. These spillovers imply that the same firm faces different labor supply curves in the pre- and post-periods.

I propose to recover the elasticity $\boldsymbol{\varepsilon}_{j\tau_1}(\chi)$ from a ratio of DiD estimands, which control for a worker's skill type and a firm's effective wage in the pre-period. I define this estimand below.

Definition. Let $\left(\ell_{j\tau_0}(\chi), w^{\text{eff}}_{j\tau_0}(\chi)\right)$ and $\left(\ell_{j\tau_1}(\chi), w^{\text{eff}}_{j\tau_1}(\chi)\right)$ be the log labor and effective wage for skill type χ at firm j in periods τ_0 and τ_1, respectively. The estimand $\text{DiD}_{\tau_0,\tau_1}(w|\chi)$ equals:

$$\frac{\mathrm{E}\left[\ell_{j\tau_1}(\chi) - \ell_{j\tau_0}(\chi) \middle| Z_j = 1, w^{\text{eff}}_{j\tau_0}(\chi) = w\right] - \mathrm{E}\left[\ell_{j\tau_1}(\chi) - \ell_{j\tau_0}(\chi) \middle| Z_j = 0, w^{\text{eff}}_{j\tau_0}(\chi) = w\right]}{\mathrm{E}\left[w^{\text{eff}}_{j\tau_1}(\chi) - w^{\text{eff}}_{j\tau_0}(\chi) \middle| Z_j = 1, w^{\text{eff}}_{j\tau_0}(\chi) = w\right] - \mathrm{E}\left[w^{\text{eff}}_{j\tau_1}(\chi) - w^{\text{eff}}_{j\tau_0}(\chi) \middle| Z_j = 0, w^{\text{eff}}_{j\tau_0}(\chi) = w\right]}.$$

The numerator of this expression is a DiD estimand corresponding to changes in log labor at a firm, while the denominator is a DiD estimand for changes in effective log earnings at a firm.

[21] A firm's effective wages can be recovered in logs as follows: $\log W^{\text{eff}}_{j\tau}(\chi) = \mathrm{E}\left[\log W_{j\tau}(\chi, \varphi) \mid j, \chi, \tau\right]$.

To show that this estimand corresponds to the labor supply elasticity, I first prove that the common trends assumption holds. Specifically, I show that—after controlling for a firm j's initial effective wage $w^{\text{eff}}_{j\tau_0}(\chi)$—the change in untreated potential outcomes over time, denoted by $w^{\text{eff}}_{j\tau_1,0}(\chi) - w^{\text{eff}}_{j\tau_0,0}(\chi)$ and $\ell_{j\tau_1,0}(\chi) - \ell_{j\tau_0,0}(\chi)$, can be written as a deterministic function of the firm-level structural parameters Γ_j, where $\Gamma_j = \left(\rho_j, \alpha_j, \{\theta_{j\chi}\}_\chi, \{a_j(X) - a_j(X')\}_{X,X'}\right)'$. Assumption I ensures that a firm j's treatment status Z_j does not depend on the parameters Γ_j. Therefore, under this assumption, treated and untreated firms with the same initial effective wages would, on average, exhibit the same change in their untreated potential outcomes.

Proposition 2. Suppose that Assumptions I and II hold. Then, given a firm's effective wage at τ_0, the change in untreated potential outcomes is mean independent of the treatment status:

$$\mathrm{E}\left[\ell_{j\tau_1,0}(\chi) - \ell_{j\tau_0,0}(\chi) \middle| Z_j = 1, w^{\text{eff}}_{j\tau_0}(\chi) = w\right] = \mathrm{E}\left[\ell_{j\tau_1,0}(\chi) - \ell_{j\tau_0,0}(\chi) \middle| Z_j = 0, w^{\text{eff}}_{j\tau_0}(\chi) = w\right].$$

$$\mathrm{E}\left[w^{\text{eff}}_{j\tau_1,0}(\chi) - w^{\text{eff}}_{j\tau_0,0}(\chi) \middle| Z_j = 1, w^{\text{eff}}_{j\tau_0}(\chi) = w\right] = \mathrm{E}\left[w^{\text{eff}}_{j\tau_1,0}(\chi) - w^{\text{eff}}_{j\tau_0,0}(\chi) \middle| Z_j = 0, w^{\text{eff}}_{j\tau_0}(\chi) = w\right].$$

Proposition 2 is central to my analysis. It implies that, by using a DiD research design, I can compare the treated and untreated potential outcomes for a firm in the post-period, thereby capturing the comparative statics from the ideal experiment. By this result, I demonstrate that the elasticity $\boldsymbol{\varepsilon}_{j,\tau_1}(\chi)$, as defined in equation (25), is point identified from the estimand $\text{DiD}_{\tau_0,\tau_1}\left(w^{\text{eff}}_{j\tau_0}(\chi)|\chi\right)$ for any firm j and skill type χ. Moreover, by evaluating these estimands over the effective wage distribution, I can recover the labor supply elasticities faced by each firm in the economy. As I discuss in Section V, this task relies on an *overlap condition*, asserting that $\mathrm{P}(Z_j = 1|w^{\text{eff}}_{j\tau_0}(\chi) = w) \in (0, 1)$ for all w in the support of $w^{\text{eff}}_{j\tau_0}(\chi)$. This condition guarantees that there are treated and untreated firms across the effective wage distribution.

Sources of Instruments

Up to this point, I have not specified how to obtain the instrument Z_j. In my analysis, I consider both internal instruments—derived from existing data based on the model's implied restrictions—and external instruments, which draw on outside data sources. Both types of instruments may be used within my framework. However, they rely on different assumptions.

My main specification utilizes internal instruments, which are constructed by specifying a process by which the firm experiences productivity shocks. This approach follows Lamadon et al. (2022) by assuming that the firm's TFP evolves according to an autoregressive process, which has a unit root. Specifically, for each time period τ, I assume that log TFP is given by:

$$t_{j\tau} = \bar{t}_j + \tilde{t}_{j\tau}, \quad \text{where} \quad \tilde{t}_{j\tau} = \tilde{t}_{j\tau-1} + u_{j\tau}. \tag{26}$$

To ensure that the productivity shocks do not lead to large jumps in wages and labor, I assume that $u_{j\tau}$ is small relative to the log wage levels at firm j. I also make the following assumption.

Assumption III *(Identification Assumptions for Internal Instruments).*
III.1. *(Instrument Relevance)* $\text{Var}(u_{j\tau}) > 0$.
III.2. *(Instrument Independence)* $u_{j\tau} \perp \left(\rho_j, \alpha_j, \{\theta_{j\chi}\}_\chi, \{a_j(X) - a_j(X')\}_{X,X'}\right)$.

Assumption III.1 asserts that productivity shocks vary across firms, which will imply that the internal instrument is relevant. Assumption III.1 guarantees that the internal instrument will satisfy Assumption I, which is sufficient to ensure that the common trends assumption holds.

Given this setup, I can construct the internal instruments from differences in the log value added of firms, $\Delta y_{j\tau}$, within a given calendar year τ. For integers $p, p' \geq 2$, I define the pre-period as $\tau_0 = \tau - p$, representing p years before τ, and the post-period as $\tau_1 = \tau + p'$, representing p' years after τ. Under Assumption III, I derive the following moment condition:

$$\text{E}\left[\Delta y_{j\tau}\left[\left(\ell_{j\tau_1}(\chi) - \ell_{j\tau_0}(\chi)\right) - \text{DiD}_{\tau_0,\tau_1}(w|\chi)\left(w^{\text{eff}}_{j\tau_1}(\chi) - w^{\text{eff}}_{j\tau_0}(\chi)\right)\right] \;\middle|\; w^{\text{eff}}_{j\tau_0}(\chi) = w\right] = 0. \quad (27)$$

This moment condition reflects an IV regression of long differences in log employment on long differences in log effective wages, instrumented by short differences in log value added, while controlling for effective wages in the pre-period. When implementing this procedure, I will define treatment status Z_j by separating firms at the median of the distribution of $\Delta y_{j\tau}$.[22]

To support the analyses using internal instruments, I also consider an external instrument. This instrument is based on plausibly exogenous product demand shocks arising from public procurement auction outcomes in Norway. This supplementary analysis follows the research design used by H. de Frahan et al. (2024), where a firm is classified as treated ($Z_j = 1$) in the year τ when it wins its first public procurement contract. Since this instrument is defined from outside data sources, it does not require making assumptions about the evolution of firm TFP. However, the main limitation of this instrument is that data on auction outcomes is only available for a subset of firms in the sample and for a limited time period, from 2003 to 2018.

IV.B. Identification of Technology, Amenities, and Preferences

I now demonstrate how to recover the structural parameters characterizing firm productivity, technology, and non-wage amenities, as well as the distribution of workers' preferences. These parameters are needed to determine rents for workers and firms, and they also allow me to run counterfactual analyses exploring sources of allocative inefficiency and inequality.

Firm Productivity and Technology

I begin by proving identification of the production parameters $(T_{j\tau}, \{\theta_{j\chi}\}_\chi, \rho_j, \alpha_j)$ for

[22]In the Appendix, I show how to extend this framework to allow for transitory measurement error in value added and labor. I require that the errors are mutually independent and also independent of firm-level parameters.

an arbitrary firm j and time period τ. Consistent with my previous analysis, I assume that the firm's TFP parameter $T_{j\tau}$ can evolve over time, while the parameters $(\{\theta_{j\chi}\}_\chi, \rho_j, \alpha_j)$ remain constant. In addition, I normalize the firm-specific efficiencies $\{\theta_{j\chi}\}_\chi$ by setting $\theta_{j\chi^*} = 1$ for some skill type $\chi^* \in \mathcal{X}$, as I can only recover these terms up to scale relative to the firm's TFP.

I demonstrate that identification relies on the ability to recover the labor supply elasticities faced by the firm at two different time periods. This task is feasible in my context because I have access to a 28-year panel, enabling me to replicate the analysis from Section IV.A over multiple years. For the time periods τ and τ', respectively, I define $\boldsymbol{\varepsilon}_{j\tau}(\chi)$ and $\boldsymbol{\varepsilon}_{j\tau'}(\chi)$ as the labor supply elasticities for firm j corresponding to skill type χ. Using these elasticities, I can evaluate the marginal products of labor at the firm. Specifically, for $\tilde{\tau} \in \{\tau, \tau'\}$, I define:

$$\text{MPL}^{\text{eff}}_{j\tilde{\tau}}(\chi) = \frac{1 + \boldsymbol{\varepsilon}_{j\tilde{\tau}}(\chi)}{\boldsymbol{\varepsilon}_{j\tilde{\tau}}(\chi)} \times W^{\text{eff}}_{j\tilde{\tau}}(\chi) \tag{28}$$

as the effective marginal product for skill type χ at firm j, adjusting for workers' skill levels φ.

Using the structural wage equation (10), I can recover the substitution parameter ρ_j by analyzing how the relative effective marginal products and effective labor shares of different skill types change over time at a firm j. This result is formalized by the following proposition.

Proposition 3. For each firm j, the elasticity of substitution between different skill types is:

$$(1 - \rho_j)^{-1} = \frac{\log\left(\frac{L^{\text{eff}}_{j\tau}(\chi)}{L^{\text{eff}}_{j\tau}(\chi')}\right) - \log\left(\frac{L^{\text{eff}}_{j\tau'}(\chi)}{L^{\text{eff}}_{j\tau'}(\chi')}\right)}{\log\left(\frac{\text{MPL}^{\text{eff}}_{j\tau}(\chi)}{\text{MPL}^{\text{eff}}_{j\tau}(\chi')}\right) - \log\left(\frac{\text{MPL}^{\text{eff}}_{j\tau'}(\chi)}{\text{MPL}^{\text{eff}}_{j\tau'}(\chi')}\right)}, \quad \text{for any } \chi, \chi' \in \mathcal{X}.$$

Proposition 3 provides a new expression for the elasticity of substitution in terms of identified quantities in the model. This expression reflects how the relative demands for two labor inputs depend on their relative productivities, as measured by the elasticity of $L^{\text{eff}}_{j\tau}(\chi)/L^{\text{eff}}_{j\tau}(\chi')$ with respect to $\text{MPL}^{\text{eff}}_{j\tau}(\chi)/\text{MPL}^{\text{eff}}_{j\tau}(\chi')$. In my framework, this relationship is specific to each firm.

Once the substitution parameter ρ_j is recovered, I prove that $\{\theta_{j\chi}\}_\chi$, α_j, and T_j can be computed sequentially from the effective marginal product, labor, and value added at the firm.

Proposition 4. For any firm j, the parameters $\{\theta_{j\chi}\}_{\chi}$, α_j, and T_j are defined by the equations:

$$\theta_{j\chi} = \exp\left[\log\left(\frac{\text{MPL}^{\text{eff}}_{j\tau}(\chi)}{\text{MPL}^{\text{eff}}_{j\tau}(\chi^*)}\right) + (1-\rho_j)\log\left(\frac{L^{\text{eff}}_{j\tau}(\chi)}{L^{\text{eff}}_{j\tau}(\chi^*)}\right)\right].$$

$$1-\alpha_j = \exp\left[\log \text{MPL}^{\text{eff}}_{j\tau}(\chi) - y_{j\tau} - \log(\theta_{j\chi}) + (1-\rho_j)\log L^{\text{eff}}_{j\tau}(\chi) + \log\sum_{\chi'\in\mathcal{X}}\theta_{j\chi'}\left(L^{\text{eff}}_{j\tau}(\chi')\right)^{\rho_j}\right].$$

$$T_{j\tau} = \exp\left[y_{j\tau} - \frac{1-\alpha_j}{\rho_j}\log\sum_{\chi\in\mathcal{X}}\theta_{j\chi}\left(L^{\text{eff}}_{j\tau}(\chi)\right)^{\rho_j}\right].$$

Note that recovering the substitution parameter ρ_j is the only step that requires knowledge of labor supply elasticities for two different time periods. If this parameter is already identified, or if one assumes that there is perfect substitutability between skill types (i.e., $\rho_j = 1$), then the other parameters $\{\theta_{j\chi}\}_{\chi}$, α_j, and T_j can be recovered from a single time period of data.

Firm Amenities

Next, I prove identification of the parameters $\{a_{j\chi}\}_{j,\chi}$, representing the average amenities at each firm j for any skill type χ. In my analysis, these terms are only identified up to scale, as they can be shifted in a way that does not impact workers' labor supply decisions. Therefore, I will normalize the amenities by defining a reference firm j^* such that $a_{j^*\chi} = 0$ for all $\chi \in \mathcal{X}$.

To recover firms' amenities, I use a revealed preference approach. This method is based on the premise that, after accounting for the differences in workers' wages, a firm that has superior amenities for a given skill type χ would attract more workers of that type. Consequently, by comparing the employment levels among firms in a counterfactual setting where all firms offer the same wages, I could draw inference about the differences in firms' amenities.

To make these comparisons, I utilize the labor supply elasticity curves $\boldsymbol{\varepsilon}_{j\tau}(\chi, w)$, which are identified for each skill type χ by the estimand $\text{DiD}_{\tau_0,\tau_1}(w|\chi)$. By integrating along these curves, I can infer the labor supply to any firm j if it were to offer a different wage, holding fixed all other wages in the economy. In doing so, I can evaluate the amenities at firm j by comparing j's labor supply to that of the reference firm j^* in the case where j and j^* pay the same wages. To illustrate this identification strategy, consider the following proposition.

Proposition 5. Under Assumptions I and II, the average amenity for skill type χ at firm j is:

$$a_{j\chi} = \ell_{j\tau}(\chi) + \int_{w^{\text{eff}}_{j\tau}(\chi)}^{w^{\text{eff}}_{j^*\tau}(\chi)} \boldsymbol{\varepsilon}_{j\tau}(\chi, w)dw - \ell_{j^*\tau}(\chi).$$

Worker Preferences

I conclude this section by establishing identification of the densities $\{f_{\beta|\chi,\tau}\}_\chi$ that characterize the distribution of workers' marginal rates of substitution between wages and amenities. This identification is achieved through revealed preference, establishing a direct link between the distribution of workers' preferences and the shape of the labor supply curve faced by a firm.

To understand the source of identification, consider how the distribution of β influences a firm's labor supply curve. When β is homogeneous, this curve is isoelastic, and the elasticity of labor supply is equal to β. However, when β is heterogeneous, the curve is nonlinear, and the labor supply elasticity depends on the firm's wage. As shown by equations (13) and (14), the nature of this dependence carries information about the moments of the distribution of β.

In the following proposition, I demonstrate that this relationship extends even further. In particular, the elasticity curve $\boldsymbol{\varepsilon}_{j\tau}(\chi, w)$ uniquely corresponds to the density $f_{\beta|\chi,\tau}(\beta|\chi,\tau)$.

Proposition 6. Given wages and labor $\{W^{\text{eff}}_{j\tau}(\chi), L_{j\tau}(\chi)\}_{j=1}^{J}$, there is a one-to-one mapping between the labor supply elasticity curve $\boldsymbol{\varepsilon}_{j\tau}(\chi, w)$ and the density function $f_{\beta|\chi,\tau}(\beta|\chi,\tau)$.

Proposition 6 implies that the density $f_{\beta|\chi,\tau}$ is point identified under same the assumptions that are needed to recover the labor supply elasticity curve $\boldsymbol{\varepsilon}_{j\tau}(\chi, w)$. The proof leverages the fact that a firm's labor supply curve, as characterized by (12), can be expressed as a Laplace transform. Moreover, I show that the inverse of this transform is proportional to the density of β. This property establishes the invertibility of a firm's labor supply curve to obtain $f_{\beta|\chi,\tau}$.[23]

V. Estimation Strategy and Empirical Results

I now describe the estimation procedure and present the empirical results, which include estimates of firm-specific labor supply elasticities and structural parameters. My results show significant heterogeneity in the labor supply elasticities, indicating that wage markdowns for workers vary considerably both within and across firms. Additionally, I find that firms offering lower wages to workers in a given skill group are more likely to face smaller elasticities. This pattern aligns with the model's predictions, suggesting that firms with lower effective wages tend to exert greater wage-setting power. When comparing across skill groups, I find that the labor supply to firms is generally less elastic for more experienced workers. Finally, my estimates indicate that there is substantial variation in firm productivity, technology, and non-wage amenities, suggesting that these factors all contribute to systematic worker sorting.

V.A. Estimation of Firm-Specific Labor Supply Elasticities

This subsection addresses the estimation of labor supply elasticities. I begin by describing my methodology and specification choices, followed by a discussion of the empirical findings.

[23] Any two functions share the same Laplace transform only if they differ on a set of Lebesgue measure zero.

Estimation Procedure and Specifications

My estimation procedure is based on the identification strategy outlined in Section IV.A. This approach consists of two key steps. First, I estimate workers' skill levels by implementing a two-way fixed effects estimator, regressing log wages on a worker fixed effect γ_i and a fixed effect $\psi_{\chi(i,\tau),j(i,\tau),\tau}$ for the skill type, firm, and year. From this regression, I define a worker's skill level φ_i as $\exp(\gamma_i)$ and the effective wage $W^{\text{eff}}_{j\chi\tau}$ as $\exp(\psi_{j(i,\tau),\chi(i,\tau),\tau})$. Table 2 presents basic summary statistics on workers' skill levels, with more details provided in the Appendix.

Next, I estimate the labor supply elasticities for each firm by computing a sample analogue of the estimand $\text{DiD}_{\tau_0,\tau_1}(w|\chi)$. This estimand corresponds to a ratio of DiD estimands, controlling for a firm's effective wage in the pre-period. Importantly, because there is a finite number of firms in the data, it is unlikely that I would observe multiple firms offering exactly the same effective wages. To overcome this issue, I employ a nonparametric Kernel estimator:

$$\frac{\sum_j K_{1,j}(w)\mathbf{1}\{Z_j=1\}\big(\ell_{j\tau_1}(\chi)-\ell_{j\tau_0}(\chi)\big)-\sum_j K_{0,j}(w)\mathbf{1}\{Z_j=0\}\big(\ell_{j\tau_1}(\chi)-\ell_{j\tau_0}(\chi)\big)}{\sum_j K_{1,j}(w)\mathbf{1}\{Z_j=1\}\big(w^{\text{eff}}_{j\tau_1}(\chi)-w^{\text{eff}}_{j\tau_0}(\chi)\big)-\sum_j K_{0,j}(w)\mathbf{1}\{Z_j=0\}\big(w^{\text{eff}}_{j\tau_1}(\chi)-w^{\text{eff}}_{j\tau_0}(\chi)\big)}, \quad (29)$$

where $K_{1,j}(w)$ and $K_{0,j}(w)$ denote the Kernel weights corresponding to treated and untreated populations, respectively. This estimator allows me to compare the outcomes of firms with similar, though not identical, initial wages, thereby ensuring that the method is tractable for finite sample sizes. In Appendix C.6, I prove that this estimator is consistent. Additionally, I derive the asymptotic properties of the estimator, demonstrating that it is capable of inference.

In my baseline specification, I utilize a Gaussian kernel function to construct the weights $K_{z,j}(w)$.[24] As I show in the Appendix, the point estimates are robust to alternative kernel functions, including the Uniform kernel, which corresponds to a standard nonparametric binning estimator. I also investigate the sensitivity of the estimates to the bandwidth choice, finding that the results remain relatively stable across varying bandwidths. Collectively, these robustness analyses provide confidence in the estimator's ability to approximate $\text{DiD}_{\tau_0,\tau_1}(w|\chi)$.

My primary estimates are based on internal instruments, following the same identification arguments laid out in Section IV.A. For each year τ that TFP shocks occur, I define treatment status Z_j by separating firms at the median of the distribution of $\Delta y_{j\tau}$. I then define the pre- and post-periods to be $\tau_0 = \tau - 2$ and $\tau_1 = \tau + 3$, respectively. I estimate the model repeatedly using all years of available data, with the post-period τ_1 spanning from 2002 to 2017. As a robustness check, I also calibrate the model using estimates obtained with external instruments based on public procurement contract wins in Norway. These calibration results draw on estimates from H. de Frahan et al. (2024), who analyze the same Norwegian admin-

[24] Specifically, I set $K_{z,j}(w) = \frac{1}{\sqrt{2\pi}}\exp\big[-\frac{1}{2}\big(\frac{1}{h}(w^{\text{eff}}_{j\tau_0}(\chi)-w)\big)^2\big]$ where the parameter h governs the bandwidth.

istrative dataset. I present my findings in the Appendix, along with a description of the data sources used in the original paper. Both internal and external instruments yield similar results.

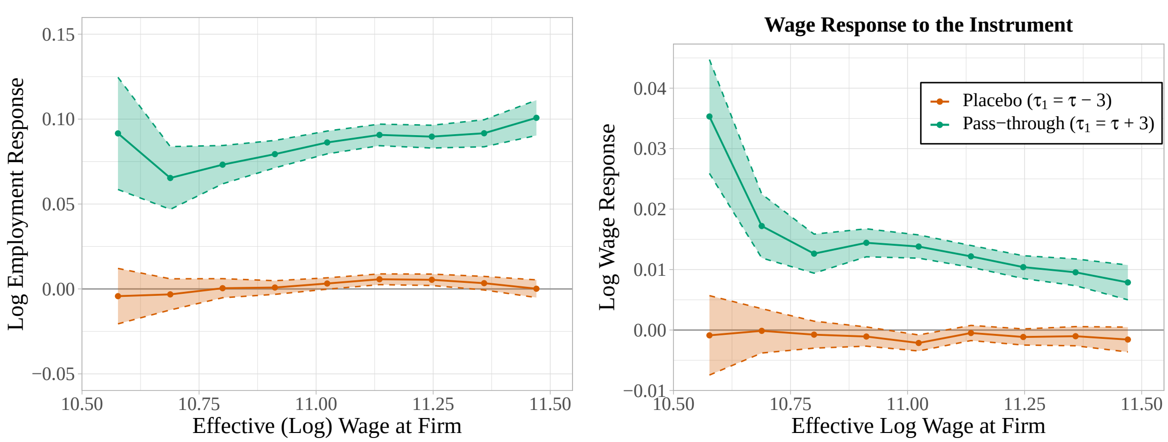
When implementing my approach, I include fixed effects for local labor markets, which represent the interactions between industries and commuting zones as defined in Section III. For comparison, I also estimate the model without fixed effects, treating Norway as one single labor market. Those results can be found in the Appendix. I average my estimates across all available years of data. Since my framework allows the labor supply elasticities to vary over time, my estimates may be interpreted as an average of these quantities across multiple years.

Description of Empirical Findings

Figure 3 presents the reduced form estimates, averaged across workers' skill types and plotted over the distribution of effective wages. These estimates correspond to the numerator and denominator of the DiD estimator defined in equation (29). In this figure, the green line displays estimates for $\tau_1 = \tau + 3$, while the orange line depicts the difference in pre-period trends between treated and untreated firms, with τ_1 set to $\tau - 3$. Notably, the placebo effect does not significantly differ from zero, lending further support to the validity of the IV design.

Figure 3: Average Pass-Through Rates by Effective Wage

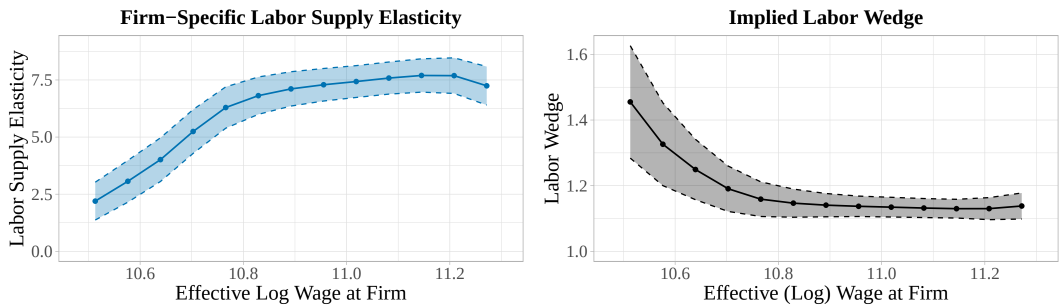
(a) Labor Response to the Instrument (b) Wage Response to the Instrument



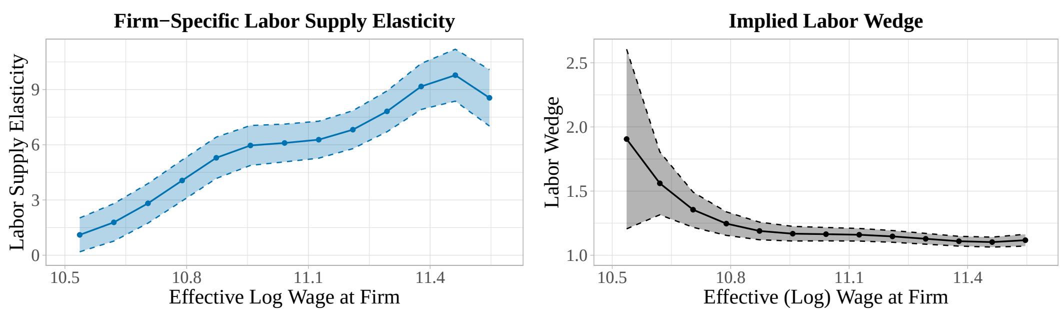
Notes. This figure plots the reduced form estimates, averaged across skill types and years $\tau_1 \in \{2002, \ldots, 2017\}$. This specification includes market fixed effects. Standard errors are bootstrapped using 500 bootstrap samples.

The reduced form estimates reveal two distinct patterns. First, the log labor response to the instrument consistently ranges between 5% and 10%, suggesting that employment growth is relatively independent of firm size. This observation aligns with Gibrat's law. Second, the log wage response ranges from 1% to 4%, and it decreases as a firm's effective wage rises. Based on equation (20), this pattern may suggest that higher-paying firms have lower returns to scale. Additionally, it could indicate that these firms face larger labor supply elasticities.

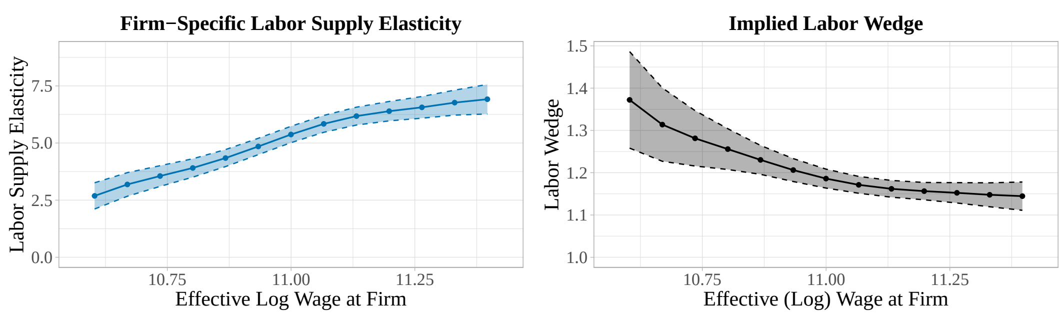
Figure 4: IV Estimates of Labor Supply Elasticities and Labor Wedges



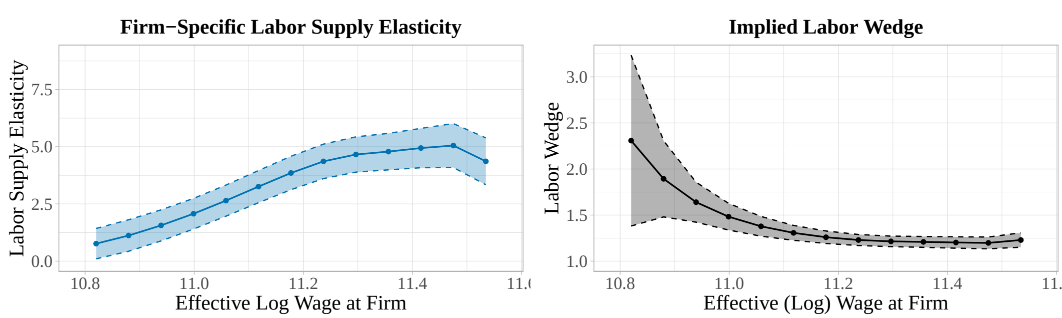
(a) No College Degree, < 10 Years Experience



(b) College Graduate, < 10 Years Experience



(c) No College Degree, ≥ 10 Years Experience



(d) College Graduate, ≥ 10 Years Experience

Notes. This figure plots IV estimates of firm-specific labor supply elasticities and labor wedges. These estimates are averaged across all years of available data, with $\tau_1 \in \{2002, \ldots, 2017\}$. The subgroup "No College Degree" represents a worker-weighted average of two different skill types: "Lower Secondary" and "Upper Secondary". This specification includes market fixed effects. Standard errors are bootstrapped using 500 bootstrap samples.

Figure 4 presents the IV estimates of firm-specific labor supply elasticities, disaggregated by workers' college attainment and experience. The figure also plots the implied labor wedge at each firm, defined as the ratio of the marginal product to a worker's wage. These estimates reveal that the labor supply elasticities differ substantially across firms, ranging between 1.5 and 10.0 for certain skill types. Moreover, they indicate that firms with higher effective wages face larger elasticities. This pattern is generally similar across worker skill types, and it implies that market power is more concentrated at the lower end of the effective wage distribution.

Although the estimated labor supply elasticities are broadly consistent across skill types, I observe notable differences based on a worker's experience. Specifically, I find that workers who have less job experience tend to exhibit larger labor supply elasticities. In addition, these elasticities are considerably more variable across firms. Under my model, these differences imply that less experienced workers are more likely to prioritize higher wages when choosing a job, and that there is also more dispersion in wage-amenity trade-offs among these workers.

To further validate my approach, I conduct two additional robustness checks. First, I test whether the instrument Z_j impacts labor on the intensive margin by influencing work hours. This test addresses a potential concern that firms respond to productivity shocks by adjusting workers' scheduled hours instead of their wages. If this were the case, then it would threaten the validity of my IV design, which assumes that the entire pass-through of demand shocks occurs in the form of wage adjustments, not changes in work hours. I find no evidence that these shocks significantly affect workers' hours, which reinforces the credibility of my assumptions.

Second, I test the assumption that firms are strategically small in the economy. A testable implication of my model is that firms offering identical wages face the same labor supply elasticities. However, this property need not hold if firms are strategically large. In a concentrated labor market, firms that offer the same wages could still face different labor supply elasticities if they have different market shares. Given this property, I test whether a firm's labor supply elasticity varies with respect to the market share, after controlling for the effective wages. I find no evidence of a market share effect, which lends support to the validity of my framework.

V.B. Estimation of Structural Parameters in the Model

Next, I present estimates of the structural parameters in the model, including the production parameters $(T_{j\tau}, \{\theta_{j\chi}\}_\chi, \rho_j, \alpha_j)$, amenities $a_{j\chi}$, and the densities $f_{\beta|\chi,\tau}$ that characterize workers' preferences. These estimates are summarized in Table 2, as well as Figures 5 and 6.

Firm Productivity, Technology, and Amenities

To estimate the production parameters, I apply the formulas from Propositions 3 and 4, which rely on the elasticity estimates, the effective wages, and the value added for each firm. I perform these calculations separately for every firm j and year τ, giving me a unique set of estimates for each combination of j and τ in the sample.[25] In Figure 5, I plot the distributions of firms' estimated TFPs $T_{j\tau}$, returns to scale $1-\alpha_j$, and substitution parameters ρ_j, averaged across all years of data. The means and variances of these quantities are reported in Table 2. I observe significant variation in firms' returns to scale and substitution parameters, suggesting that differences in technology across firms may play an important role in wage determination.

Table 2: Estimates of Structural Parameters—Means and Variances

	Notation	Mean	Variance
Panel A. Firm Production			
Log Total Factor Productivity	$t_{j\tau}$	12.92 (0.05)	1.19 (0.01)
Returns to Scale	$1-\alpha_j$	0.70 (0.01)	0.04 (0.01)
Substitution Parameter	ρ_j	0.72 (0.04)	0.54 (0.13)
Panel B. Firm Amenities	$a_{j\chi}$		
No College, < 10 Years Exp.		—	1.54 (0.06)
College, < 10 Years Exp.		—	2.02 (0.02)
No College, ≥ 10 Years Exp.		—	1.63 (0.11)
College, ≥ 10 Years Exp.		—	1.60 (0.04)
Panel C. Worker Skill Levels	φ_i		
No College, < 10 Years Exp.		—	0.062 (0.01)
College, < 10 Years Exp.		—	0.073 (0.02)
No College, ≥ 10 Years Exp.		—	0.064 (0.00)
College, ≥ 10 Years Exp.		—	0.076 (0.01)
Panel D. Worker Preferences	β_i		
No College, < 10 Years Exp.		6.30 (0.76)	13.76 (1.01)
College, < 10 Years Exp.		7.23 (1.05)	16.25 (2.71)
No College, ≥ 10 Years Exp.		5.96 (0.42)	15.71 (0.87)
College, ≥ 10 Years Exp.		4.28 (0.68)	15.11 (2.25)

Notes. This table presents estimates of cross-sectional means and variances of model parameters, averaged across years. The means are omitted for parameters that are only identified up-to-scale.

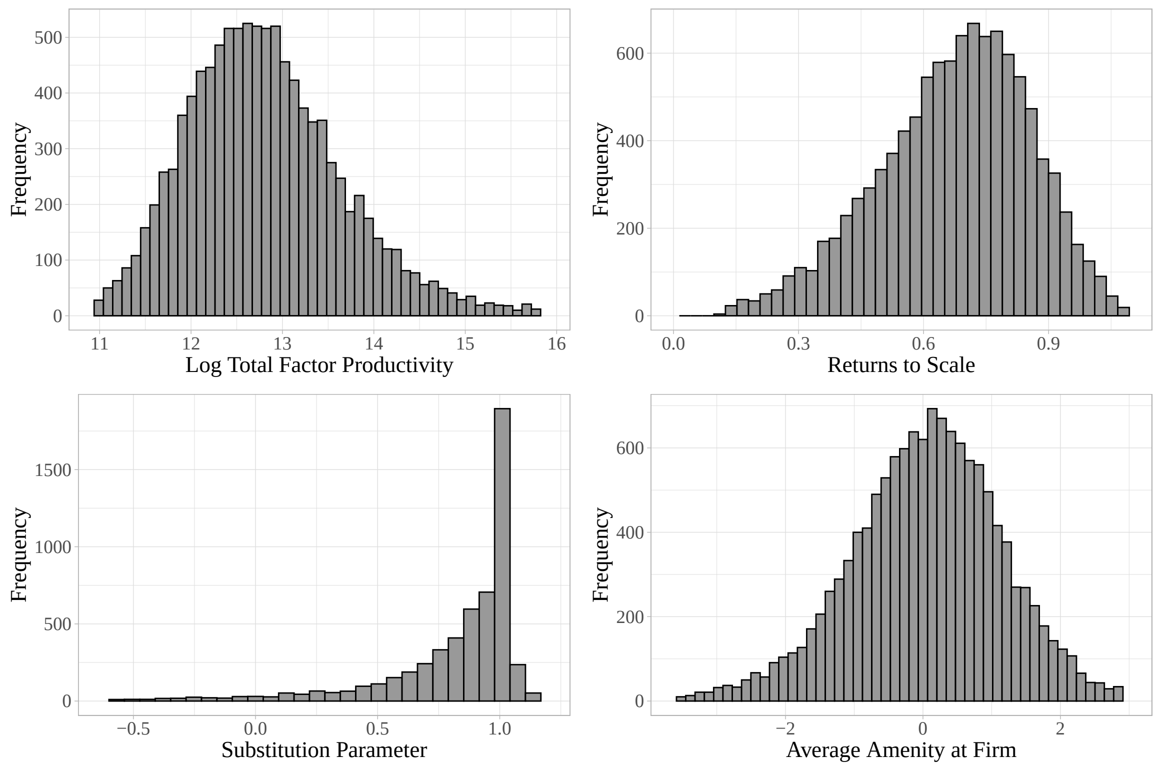
To estimate non-wage amenities, I follow the procedure laid out in Proposition 5, which involves integrating the estimated labor supply elasticities with respect to a firm's effective wage.[26] This approach allows me to compare labor across firms in a counterfactual setting where firms offer identical wages. I plot the estimated distribution of firms' average amenities in Figure 5, with variances by subtype shown in Table 2. I find that the amount of variation

[25] Recall that computing the substitution parameter ρ_j for a firm requires me to compare estimates over two different time periods: τ and τ'. In my implementation, I define τ' as $\tau - 1$, corresponding to the previous year.

[26] For each skill type χ, amenities are normalized relative to a "reference firm" j^*, where $a_{j^*\chi} = 0$. In my analysis, j^* is defined as the firm offering an average wage equal to the median of the distribution in the sample.

in amenities is generally similar across worker types. However, for less experienced workers, it appears that those with a college education face greater variation in amenities across firms.

Figure 5: Histograms of Estimated Firm Productivity, Technology, and Amenities



Notes. This figure plots histograms of the estimated (log) total factor productivity $\log T_{j\tau}$, the returns to scale $1-\alpha_j$, the substitution parameter ρ_j, and the average amenities $\mathrm{E}(a_{j\chi}|j)$ across firms in the estimation sample.

Worker Preferences

To estimate the density function $f_{\beta|\chi,\tau}$, I draw on Proposition 6. This result establishes that this density is uniquely determined by the labor supply curve faced by an individual firm. Consequently, $f_{\beta|\chi,\tau}$ may be recovered as the unique solution to the following equations:

$$L_{j\tau}(\chi, W) = \int \frac{\exp\left(\beta \log W + a_{j\chi}\right)}{I_\tau^{\text{eff}}(\beta, \chi)} f_{\beta|\chi,\tau}(\beta|\chi,\tau) f_{\chi|\tau}(\chi|\tau) d\beta, \quad \text{for } W > 0. \tag{30}$$

Note that the labor supply curve $L_{j\tau}(\chi, W)$, the effective wage indices, $I_\tau^{\text{eff}}(\beta, \chi)$, the density of skill types $f_{\chi|\tau}(\chi|\tau)$, and the non-wage amenities $a_{j\chi}$ can all be recovered from the data.

To make the estimation problem tractable, I introduce two important modifications. First, given that the density function $f_{\beta|\chi,\tau}$ is an infinite-dimensional parameter, I will approximate it using a finite-dimensional basis expansion. Specifically, I replace $f_{\beta|\chi,\tau}$ with $\hat{f}_{\beta|\chi,\tau}$, where:

$$\hat{f}_{\beta|\chi,\tau}(b) = \sum_{k=1}^{d_\phi} \phi_k \hat{f}^{(k)}_{\beta|\chi,\tau}(b), \quad \text{for } \phi = (\phi_1, \ldots, \phi_{d_\phi}) \in \Phi. \tag{31}$$

In this expression, ϕ is a d_ϕ-dimensional vector of unknown coefficients with support Φ, and $\{\hat{f}^{(k)}_{\beta|\chi,\tau}\}_k$ are basis functions known to the researcher. By defining the basis in this way, I can approximate $f_{\beta|\chi,\tau}$ with a function $\hat{f}_{\beta|\chi,\tau}$ that is parametrized by a finite-dimensional vector ϕ.

Second, rather than imposing an infinite number of constraints, I evaluate equation (30) on a finite grid of effective wages $\mathbf{W} \subset \mathbb{R}_+$, noting that identification is achieved as $|\mathbf{W}| \to \infty$. This modification allows me to recast (30) as a convex optimization problem, minimizing the ℓ^2-distance between the estimated choice probabilities and the predicted probabilities under the density $\hat{f}_{\beta|\chi,\tau}$. In particular, my estimation procedure involves choosing ϕ^* by solving:

$$\phi^* = \text{argmin}_\phi \sum_{W \in \mathbf{W}} \left[\phi' \hat{h}_\tau(\chi, W) - \hat{L}_{j\tau}(\chi, W)\right]^2. \tag{32}$$

where $\hat{L}_{j\tau}(\chi, W)$ is the estimated labor supply for $W^{\text{eff}}_{j\tau}(\chi) = W$ and $\hat{h}_\tau(\chi, W) \in \mathbb{R}^{d_\phi}$ is a vector with entries $\hat{h}_{\tau,k}(\chi, W) = \int \left[\exp(\beta \log W + \hat{a}_{j\chi}) / \hat{I}^{\text{eff}}_\tau(\beta, \chi)\right] \hat{f}^{(k)}_{\beta|\chi,\tau}(\beta|\chi, \tau) \hat{f}_{\chi|\tau}(\chi|\tau) d\beta$.[27]

In my baseline specification, I construct a basis using mixtures of $d_\phi = 18$ Gamma density functions. This type of density have been shown to be particularly effective for nonparametric estimation of functions that have positive support; e.g., see Wiper et al. (2001) for discussion. As a robustness check, I investigate the sensitivity of my resulting estimates to the choice of basis functions, the basis size d_ϕ, and the number of points $|\mathbf{W}|$ in the wage grid. I find that the estimates remain robust to variation in each of these elements. Moreover, I demonstrate that the earnings and employment distributions predicted by the estimated densities $\hat{f}_{\beta|\chi,\tau}$ closely match those in the sample, which suggests that my estimates are a good fit for the data.

Figure 6 plots the estimated density of β among workers in the population. I also compare moments of this distribution for different worker types in Table 2. I find that more experienced workers tend to have lower values of β, which is associated with a greater willingness to pay for amenities. However, there is substantial variation in preferences within each worker type.

VI. Implications for Rent Sharing, Misallocation, and Inequality

I now summarize the main empirical findings from my estimates, quantifying imperfect competition in the economy and assessing its implications for welfare and earnings inequality.

[27]Note that equation (32) may be written in matrix form as $\phi^* = \text{argmin}_\phi \left\{\phi^\top \hat{h}^\top \hat{h} \phi - 2\phi^\top \hat{h}^\top \hat{L} + \hat{L}^\top \hat{L}\right\}$, where $\hat{L} \in \mathbb{R}^{|\mathbf{W}|}$ and $\hat{h} \in \mathbb{R}^{|\mathbf{W}| \times d_\phi}$. This framing of the problem is more convenient for computational implementation.

Figure 6: Density of Workers' Individual Labor Supply Elasticities

Notes. This figure plots the estimated density f_β, which corresponds to the distribution of worker's preferences over wages and amenities. The confidence region is bootstrapped using estimates from 500 bootstrap samples.

VI.A. Quantifying Imperfect Competition

Table 3 presents estimates of wage markdowns, along with worker and firm rents. On average, I find that workers are paid around 86% of their marginal product. However, these markdowns vary both within and across firms. I find that approximately 70% of the variation in markdowns occurs within firms, while the remaining 30% is due to differences across firms.

Table 3: Estimates of Wage Markdowns and Rents

	Mean	Std. Dev.	Decomposition of Variation	
			Within Firm	Between Firm
Wage Markdown	0.864 (0.011)	0.035 (0.025)	69.4%	30.6%
Worker Rents				
Per-worker Dollars	$13,469 (788)	$2,859 (2,662)	70.5%	29.5%
Share of Earnings	17.1% (0.9)			
Firm Rents				
Per-worker Dollars	$9,403 (1,451)	$4,763 (3,370)	0.0%	100.0%
Share of Profits	29.2% (2.6)			

Notes. This table presents estimates of wage markdowns and rents. Variance decompositions are calculated via the Law of Total Variance. Standard errors, shown in parentheses, are bootstrapped using 500 bootstrap samples.

I find that workers and firms capture substantial rents, with worker rents being larger, on average, than firm rents. I estimate that the average worker is willing to pay approximately \$13,500—or 17% of their earnings—to avoid working for a different firm. In contrast, the average firm obtains \$9,400 per worker—or 29% of its profit—by exerting its wage-setting power. Figure A.10 of the Appendix plots the cross-sectional densities of estimated worker and firm rents, averaged across years. I find that these rents vary considerably in the economy.

VI.B. Misallocation and Inequality

To evaluate the welfare impact of imperfect competition, I compare equilibrium outcomes in the monopsonistic economy to counterfactual outcomes that would arise in a competitive (Walrasian) economy, where firms do not exert any wage-setting power. In my framework, misallocation occurs because firms vary in their wage-setting power. This variation can lead workers to sort into firms that they would not have chosen in a competitive labor market.

Table 4: Estimates of Counterfactual Comparisons

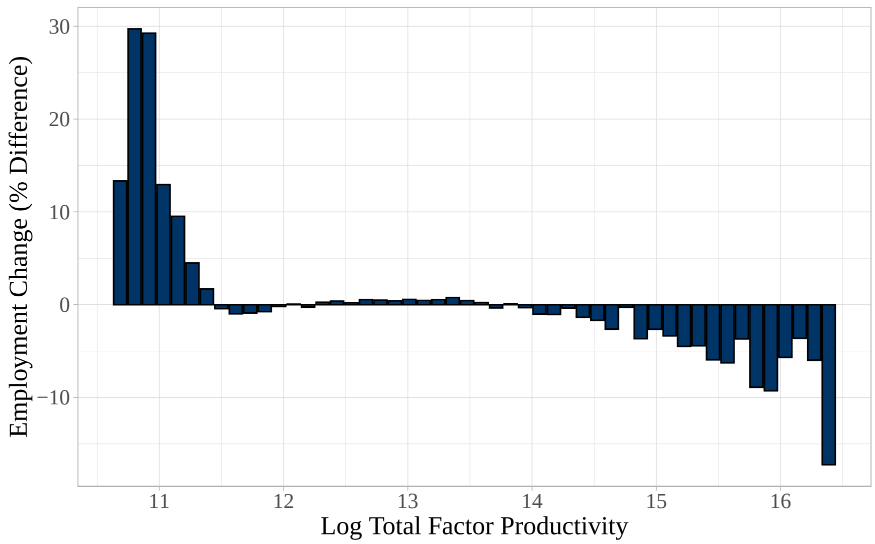
	Monopsonistic Labor Market	Competitive Labor Market	% Difference in Dollars
Total Welfare (log dollars)	12.71 (0.21)	12.81 (0.23)	9.58%
Mean Earnings (log dollars)	11.29 (0.00)	11.46 (0.01)	15.23%
Variance in Log Earnings			
Total	0.103 (0.00)	0.105 (0.01)	-2.59%
Between Firms	0.025 (0.00)	0.026 (0.01)	-1.49%
Within Firms	0.078 (0.00)	0.079 (0.00)	-0.10%

Notes. This table compares outcomes in the monopsonistic economy to counterfactual outcomes in a competitive (Walrasian) economy. The standard errors, shown in parentheses, are bootstrapped using 500 bootstrap samples.

Table 4 presents estimates for counterfactual comparisons. These estimates suggest that imperfect competition produces large misallocation of workers to firms, resulting in a 9.58% reduction in total welfare in dollar terms. Furthermore, in a competitive labor market, average earnings would increase by 15.23%. However, despite these changes, the overall impact on wage inequality, as measured by the total variance in log earnings, appears to be minimal.

To understand how workers would be reallocated in a competitive labor market, recall that I estimate larger markdowns for firms that have lower effective wages. This suggests that, relative to the competitive benchmark, workers are being allocated away from less productive firms. Indeed, I find that the correlation between a firm's log labor $\ell_{j\tau}$ and its log TFP $t_{j\tau}$ is higher (0.155) in a monopsonistic market than it would be in a competitive market (0.089). Moreover, in Figure 6, I plot the changes in estimated labor shares that would occur if labor wedges were eliminated across the distribution of firm TFP. This figure illustrates that, under the competitive allocation, more workers would be employed at firms with lower productivities. Meanwhile, fewer workers would be employed at firms with higher productivities.

Figure 6: Estimated Reallocation of Labor by Eliminating Labor Wedges



Notes. This figure depicts the reallocation of labor that results from eliminating labor wedges. It plots the estimated percent change in employment across different groups of firms, organized by their estimated log TFPs.

VII. Conclusions

This paper explores how an employer's wage-setting power depends on the characteristics of its workforce and assesses the consequences of this relationship for welfare and inequality. To this end, I develop a structural model that integrates labor market frictions with individual heterogeneity in workers' trade-offs between wages and non-wage job characteristics. I show that these two factors, taken together, cause employers to face varying labor supply elasticities, resulting in differences in labor market power both within and across firms. Moreover, I show that this framework is empirically tractable and can be used to draw inferences about sorting, imperfect competition, and rent sharing. I estimate the model using a matched worker-firm panel dataset that covers the universe of workers and firms in Norway between 1995 and 2018.

One key empirical finding from my estimates is that, after accounting for workers' skills, firms with lower wages tend to have more labor market power. My model offers an explanation for this finding. That is, when workers make different trade-offs over job attributes, they tend to sort accordingly, with workers who prioritize amenities (e.g., flexible work hours, fewer physical demands, and/or more societal impact) choosing jobs that pay less. In a labor market that is imperfectly competitive, these sorting patterns lead lower-paying firms to attract labor that is less elastic (i.e., less sensitive to wage changes), thereby increasing wage-setting power. This phenomenon has important welfare implications, because the disparities in market power between firms generate misallocation of labor relative to the competitive benchmark. In a perfectly competitive labor market, more individuals would be employed at low-wage firms.

While this paper offers an initial examination of the relationship between workers' wage-amenity trade-offs and labor market power, there is much room for future work. For example, my framework does not consider the possibility that firms strategically compete for workers by internalizing their market shares. I justify this assumption by testing whether a firm's market share impacts its labor supply elasticities after controlling for wages, finding no significant effect. However, this test does not definitively rule out the existence of strategic interactions. A meaningful extension of my framework would be to examine how a firm's wage-setting power is affected not only by the composition of its workforce but also by its market share.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67, 251–333.

Andrews, M. J., L. Gill, T. Schank, and R. Upward (2008): "High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias?" *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 171, 673–697.

Azar, J. A., S. T. Berry, and I. Marinescu (2022): "Estimating Labor Market Power," Working Paper 30365, National Bureau of Economic Research.

Azar, J., and I. Marinescu (2024): "Monopsony Power in the Labor Market: From Theory to Policy," *Annual Review of Economics*, 16, 491–518.

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

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63, 841–890.

Bhaskar, V., A. Manning, and T. To (2002): "Oligopsony and Monopsonistic Competition in Labor Markets," *Journal of Economic Perspectives*, 16, 155–174.

Bhuller, M. S. (2009): "Inndeling av Norge i Arbeidsmarkedsregioner," *Statistisk Sentralbyrås Notater*, 1–30.

Boal, W. M., and M. R. Ransom (1997): "Monopsony in the Labor Market," *Journal of Economic Literature*, 35, 86–112.

Bruun, M., J. Delbas, H. Ottosson, and kees Zeelenberg (2021): "Peer Review Report on Compliance with the European Statistics Code of Practice and Further Improvement and Development of the National Statistical System: Norway," December.

Burdett, K., and D. T. Mortensen (1998): "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, 39, 257–273.

Card, D. (2022): "Who Set Your Wage?" *American Economic Review*, 112, 1075–90.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018): "Firms and Labor Market Inequality: Evidence and Some Theory," *Journal of Labor Economics*, 36, S13–S70.

Card, D., J. Heining, and P. Kline (2013): "Workplace Heterogeneity and the Rise of West German Wage Inequality*," *The Quarterly Journal of Economics*, 128, 967–1015.

Chan, M., K. Kroft, E. Mattana, and I. Mourifié (2024): "An Empirical Framework For Matching With Imperfect Competition," Working Paper 32493, National Bureau of Economic Research.

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

Eeckhout, J., and P. Kircher (2011): "Identifying Sorting—In Theory," *The Review of Economic Studies*, 78, 872–906.

Fagereng, A., M. Mogstad, and M. Rønning (2021): "Why Do Wealthy Parents Have Wealthy Children?" *Journal of Political Economy*, 129, 703–756.

H. de Frahan, L., T. Lamadon, and T. Meling (2024): "Why Do Larger Firms Have Lower Labor Shares?".

Gibbons, R., and L. Katz (1992): "Does Unmeasured Ability Explain Inter-Industry Wage Differentials?," *The Review of Economic Studies*, 59, 515–535.

Gibbons, R., L. F. Katz, T. Lemieux, and D. Parent (2005): "Comparative Advantage, Learning, and Sectoral Wage Determination," *Journal of Labor Economics*, 23, 681–724.

Hall, R. E., and A. I. Mueller (2018): "Wage Dispersion and Search Behavior: The Importance of Nonwage Job Values," *Journal of Political Economy*, 126, pp. 1594–1637.

Hamermesh, D. S. (1999): "Changing Inequality in Markets for Workplace Amenities," *The Quarterly Journal of Economics*, 114, 1085–1123.

Jarosch, G., J. S. Nimczik, and I. Sorkin (2024): "Granular Search, Market Structure, and Wages," *The Review of Economic Studies*, rdae004.

Kroft, K., Y. Luo, M. Mogstad, and B. Setzler (2021): "Imperfect Competition and Rents in Labor and Product Markets: The Case of the Construction Industry," working paper, University of Toronto, Department of Economics.

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.

Lise, J., and F. Postel-Vinay (2020): "Multidimensional Skills, Sorting, and Human Capital Accumulation," *American Economic Review*, 110, 2328–76.

Maestas, N., K. J. Mullen, D. Powell, T. von Wachter, and J. B. Wenger (2017): "Working Conditions in the United States: Results of the 2015 American Working Conditions Survey,"Technical report, Santa Monica, CA: RAND Corporation.

——— (2023): "The Value of Working Conditions in the United States and the Implications for the Structure of Wages," *American Economic Review*, 113, 2007–47.

Manning, A. (2010): "The plant size-place effect: agglomeration and monopsony in labour markets," *Journal of Economic Geography*, 10, 717–744.

——— (2011): "Chapter 11 - Imperfect Competition in the Labor Market," Volume 4 of Handbook of Labor Economics: Elsevier, 973–1041.

——— (2021): "Monopsony in Labor Markets: A Review," *ILR Review*, 74, 3–26.

Mas, A., and A. Pallais (2017): "Valuing Alternative Work Arrangements," *American Economic Review*, 107, 3722–59.

Murphy, K. M., and R. H. Topel (1990): *Efficiency Wages Reconsidered: Theory and Evidence*, 204–240, London: Palgrave Macmillan UK.

Nevo, A. (2000): "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand," *Journal of Economics & Management Strategy*, 9, 513–548.

Pierce, B. (2001): "Compensation Inequality*," *The Quarterly Journal of Economics*, 116, 1493–1525.

Robinson, J. (1933): *Journal of the Royal Statistical Society*, 97, 671–674.

Rosen, S. (1986): "Chapter 12 The theory of equalizing differences," Volume 1 of Handbook of Labor Economics: Elsevier, 641–692.

Sokolova, A., and T. Sorensen (2021): "Monopsony in Labor Markets: A Meta-Analysis," *ILR Review*, 74, 27–55.

Song, J., D. J. Price, F. Guvenen, N. Bloom, and T. von Wachter (2018): "Firming Up Inequality," *The Quarterly Journal of Economics*, 134, 1–50.

Sorkin, I. (2018): "Ranking Firms Using Revealed Preference*," *The Quarterly Journal of Economics*, 133, 1331–1393.

Taber, C., and R. Vejlin (2020): "Estimation of a Roy/Search/Compensating Differential Model of the Labor Market," *Econometrica*, 88, pp. 1031–1069.

Verboven, F. (1996): "The Nested Logit Model and Representative Consumer Theory," *Economics Letters*, 50, 57–63.

Wiper, M., D. R. Insua, and F. Ruggeri (2001): "Mixtures of Gamma Distributions with Applications," *Journal of Computational and Graphical Statistics*, 10, 440–454.

Wiswall, M., and B. Zafar (2018): "Preference for the Workplace, Investment in Human Capital, and Gender," *The Quarterly Journal of Economics*, 133, 457–507.