NBER WORKING PAPER SERIES

AN EMPIRICAL FRAMEWORK FOR MATCHING WITH IMPERFECT COMPETITION

Mons Chan
Kory Kroft
Elena Mattana
Ismael Mourifié

Working Paper 32493
http://www.nber.org/papers/w32493

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2024

The first version is April 2nd, 2019. This present version is May 15, 2024. We thank Victor Aguirregabiria, Karen Bernhardt-Walther, Stéphane Bonhomme, Antoine Djogbenou, Yao Luo, Thibault Lamadon, Magne Mogstad, Suresh Naidu, Marcin Peski, and Aloysius Siow for helpful discussions, and seminar audiences at Western University, University of Toronto, WUSTL, UCL, Berkeley, Columbia, Wisconsin Madison, Virginia, UBC, Princeton, University of Bergen, Aarhus University, Queen's University, the US Census, University of Michigan, UIUC, Notre Dame University, and University of Chicago. We thank Renato Zimmermann for excellent research assistance. The research was conducted in part when Ismael Mourifié was visiting the Becker Friedman Institute (BFI) at the University of Chicago. Mourifié thanks his hosts for its hospitality and support. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

An Empirical Framework For Matching With Imperfect Competition
Mons Chan, Kory Kroft, Elena Mattana, and Ismael Mourifié
NBER Working Paper No. 32493
May 2024
JEL No. J0,J20,J3,J31,J42

ABSTRACT

This paper builds, identifies and estimates a model of the labor market that features strategic interactions in wage setting and two-sided heterogeneity in order to shed light on the sources of wage inequality. We provide a tractable characterization of the model equilibrium and demonstrate its existence and uniqueness. This characterization of the equilibrium allows us to derive a rich set of comparative statics and to gauge the relative contributions of worker skill, preference for amenities and strategic interactions to equilibrium wage inequality. Using instrumental variables, we establish identification of labor demand and supply parameters and estimate them using matched employer-employee data from Denmark. Using our estimated structural model, we perform a series of counterfactual analyses in order to provide a quantitative evaluation of the main sources of wage inequality in Denmark.

Mons Chan
Queen's Unviersity
Department of Economics
94 University Ave
Kingston K7L 3N6
Canada
mons.chan@queensu.ca

Kory Kroft
Department of Economics
University of Toronto
150 St. George Street
Toronto, ON M5S 3G7
and NBER
kory.kroft@utoronto.ca

Elena Mattana
Aarhus University
Department of Economics and
Business Economics
Fuglesangs Allé 4
Aarhus V 8210
Denmark
emattana@econ.au.dk

Ismael Mourifié
Department of Economics
Washington University in St. Louis
One Brookings Drive
St. Louis, MO 63130-4899
and NBER
ismaelyacoub@gmail.com

1. Introduction

It is well-known that observationally similar workers earn different wages when working at different firms. A recent literature has investigated the role of labor market power in generating firm-specific pay premia. One strand of this literature has focused on settings that allow for rich heterogeneity across workers and firms, but has maintained the assumption that firms are "atomistic" and thus abstract from strategic interactions in wage setting between firms (Card et al., 2018; Lamadon et al., 2022).[1] Another strand of the literature has considered settings where there are strategic interactions in wage setting (Jarosch et al., 2019; Berger et al., 2022a).[2] However, papers in this strand largely abstract from skill heterogeneity across workers and thus do not capture differences in earnings due to standard competitive forces such as human capital nor do they allow for sorting based on match-specific production complementarities (Roy sorting). If there is two-sided heterogeneity and strategic interactions in wage setting, neither of these approaches is likely to provide a fully accurate characterization of the sources of wage inequality.

This paper builds, identifies and estimates a model of the labor market that features strategic interactions in wage setting ("oligopsony") and two-sided heterogeneity in order to shed light on the sources of wage inequality. We provide a tractable characterization of the model equilibrium and demonstrate existence and uniqueness. This equilibrium characterization allows us to derive a rich set of comparative statics and to use counterfactuals to gauge the relative contributions of worker skill, preference for amenities and strategic interactions for equilibrium wage inequality. We use our model to characterize the main sources of endogeneity in the labor supply and labor demand equations and show how instrumental variables (IV) approaches—that have traditionally been applied in the context of differentiated product markets—facilitate identification of the labor supply and labor demand parameters. We estimate the structural parameters using matched employee-employer data from Denmark. Using our estimated structural model, we perform a series of counterfactual analyses in order to provide a quantitative evaluation of the main sources of wage inequality in Denmark.

In Section 2, we develop a many-to-one matching model of the labor market with imperfect competition building on Rosen (1986), Boal and Ransom (1997), Bhaskar et al. (2002), Card et al. (2018), and Lamadon et al. (2022). On one side of the market are a finite number of heterogeneous firms who post wages. On the other side of the market

[1] Lamadon et al. (2022) write that if local markets are segmented by geography or location, then strategic interactions can play an important role but "identification of such interaction effects is challenging with two-sided heterogeneity".

[2] These frameworks shed light on the link between employer concentration and wages, as emphasized by Benmelech et al. (2022); Rinz (2022); Azar et al. (2022a), and Azar et al. (2022b).

are a large number of workers divided into a discrete number of types with heterogeneous skills and preferences. Worker skills are multidimensional and match-specific, meaning a given worker's productivity can vary across firms thus capturing the key features of the Roy model (Roy, 1951). Workers have preferences over wages and both deterministic and stochastic preferences over employer amenities, and choose a firm or non-employment to maximize utility.[3] The heterogeneity in preferences for amenities across firms implies that wages will reflect compensating differentials, similar to Rosen (1986). The heterogeneity in stochastic preferences for amenities along the horizontal dimension implies that firms face upward-sloping labor supply curves giving rise to market power. Given this market power, firms optimally mark down wages below the marginal revenue product of labor.

There are two main sources of labor market power in our model of imperfect competition. The first source is employer differentiation due to differences in workers' stochastic preferences for amenities. Due to imperfect information, employers cannot observe the stochastic part of workers' preferences and thus, cannot perfectly wage discriminate and extract all workers' surplus. This leads firms to mark down wages while at the same time creates rents for infra-marginal workers at the firm. The second source of labor market power is employer concentration due to the presence of a finite number of firms in the labor market. For example, when the stochastic part of worker preferences follows a Nested Logit distribution and firms set wages according to Bertrand-Nash competition, strategic interactions operate through local and aggregate wage indices. In our model, firms can internalize the impact of a wage change on the market wage indices. This contrasts with models featuring monopsonistic competition, such as Lamadon et al. (2022), where firms are "atomistic" and do not internalize their impact on these wage indices so the labor supply elasticity and hence the markdown are constant. We show in this setting that a researcher who ignores strategic interactions may overestimate the markdown.[4] In combining these two sources of labor market power, our model also allows the level of concentration to depend on worker type thus introducing a new channel through which worker heterogeneity can affect wages.

Section 3 provides a tractable equilibrium characterization of our matching model with imperfect competition without restricting the level of strategic interactions. We provide general conditions on individual preferences and firm production technology under which an equilibrium exists and is unique. For uniqueness, on the worker side we require a shape

[3] Many of the existing models of monopsony, such as the ones cited above, do not consider non-employment as an option. We show how allowing for an outside option helps with identification of the structural labor supply elasticity.

[4] Berger et al. (2022a) capture strategic interactions in their framework; however, they assume that while firms can be dominant in their "local" market, they cannot be dominant in the overall economy. This assumption also naturally leads one to overestimate the true markdown but with a lower bias than the one estimated under the "atomistic" firm assumption.

restriction on the labor supply elasticities, which trivially holds in the widely-used Nested Logit model. On the firm side, we require that production functions are additively separable in labor of different types but allow for decreasing returns to scale and imperfect substitution across labor types. In the case of non-separable production functions, we characterize a set of testable implications under which the equilibrium is unique. In addition, we establish that under the same conditions, there exist globally convergent methods (Gauss-siedel or Jacobi iteration) that allow one to solve for the unique equilibrium of the model. This has an important empirical advantage, since efficiently solving for the equilibrium allows us to perform a series of counterfactual analyses to understand how different features of our model contribute to the observed wage distribution. To the best of our knowledge, this is one of the first papers in the labor literature that considers a general equilibrium model of wage setting with imperfect competition, characterizes the equilibrium (demonstrating existence and uniqueness) and uses this characterization to solve for various counterfactual scenarios.

Next, we use the equilibrium characterization to derive a set of comparative statics. First, we show that firms' strategic interactions in wage setting amplify the pass-through effect of a firm-specific productivity shock on equilibrium wages. Intuitively, a productivity shock to one firm in the market causes other firms to post higher wages and this triggers a set of successive wage responses until a new equilibrium is reached. This implies that a researcher cannot use equilibrium wage and employment responses to firm-specific shocks to identify the slope of the labor supply curve, as pointed out by Berger et al. (2022a). We derive a sharp lower bound for an exogenous change in the total factor productivity (TFP) of a firm on wages of a given worker type. This lower bound corresponds to the change one would obtain in a model without strategic interactions. Second, we consider a firm-specific amenities shock and show that the equilibrium effect on the firm's wage is ambiguous. While an increase in amenities initially lowers wages at the firm, this causes other firms in the market to increase their wages through a competition effect and this feeds back to increase wages at the original firm triggering a succession of wage changes until an equilibrium is reached. Thus, unlike with productivity shocks, one cannot conclude how strategic interactions affects the impact of amenities shocks on wages.

In Section 4, we introduce a social welfare function and study its properties. A key result is that our framework implies a natural measure of concentration—the "generalized concentration index" (GCI)—which is a function of the generalized entropy introduced in Galichon and Salanié (2022). In particular, we establish a connection between the social welfare function and the GCI. We show that under certain conditions, increases in market concentration lower social welfare. In the case of Nested Logit preferences, we show that the GCI can be expressed as a weighted function of the "within-nest" concentration values, and a "between-nest" component. As pointed out in Maasoumi and Slottje (2003), this type of decomposability of a concentration index is very useful when there is heterogeneity across local markets as it allows one to more accurately pinpoint the main sources of concentration and examine the potentially heterogeneous impact of policy changes, such as minimum wage reforms, on particular markets as well as on overall concentration. The widely used Herfindahl-Hirschman Index (HHI) does not have this decomposability feature.

In Section 5, we consider an empirical model which imposes parametric assumptions and we establish identification. Our approach follows the dominant empirical IO paradigm by developing a theory that is tied to the market and a clear analysis of endogeneity, identification, and instruments. On the worker side, we assume Nested Logit preferences and derive the quasi-supply function following Berry (1994). This function expresses the market share for a worker type and a firm relative to the share of the outside option (non-employment) for that worker type as a function of wages at the firm and the inside share. This approach allows for identification of the labor supply parameters in the presence of oligopsony and strategic interactions by directly controlling for the unobserved market index. The remaining identification problem is that wages and the inside share are correlated with unobserved (to the econometrician but not the firm) preferences for amenities. Using an instrumental variables strategy which follows Lamadon et al. (2022), we establish identification of the labor supply parameters and construct the labor supply elasticities and deterministic preferences for amenities.

Next, we exploit firm optimization to derive an estimating equation for the relative labor demand between worker types. Due to the presence of labor market power, this equation depends on the labor supply elasticities which are known from the prior step.[5] We use this estimating equation to identify the firm-level production functions which feature heterogeneous labor inputs, flexible asymmetric substitution elasticities, match-specific labor productivity, and imperfect competition in labor markets. The identification challenge is that relative labor demand is correlated with relative unobserved labor productivities. Under the assumption that labor productivity follows a first-order auto-regressive process following Doraszelski and Jaumandreu (2018), we show that an instrumental variables strategy which uses lagged revenues, employment and wages as instruments identifies the substitution parameters. Given identification of the substitution parameters, we can then use the firm's first-order conditions to recover labor productivities up to a scale normalization. Finally, under a restriction on product market competition, we can identify the joint distribution of

[5]Intuitively, at the margin, when the firm expands production, the marginal cost of a new hire is not the wage but rather the wage scaled by the labor supply elasticity since the firm must pay more to its workers that are inframarginal.

firm-specific returns to scale and total factor productivity. Our estimation approach closely follows our identification arguments thus providing a tight link between them.

In Section 6, we describe our data and report our empirical results. We first discuss the construction and details of our dataset which is built from several Danish administrative registers. The Danish administrative registers allow us to link matched employee-employer data to establishment location and firm revenue for the private sector. Therefore, we estimate our model using data on full-time employees in the private sector. We assign individuals to 12 types based on a combination of sex, age, and education. We define local labor markets as industry-commuting zone pairings. We compute the GCI for the Danish private sector and find that roughly 14 percent of local markets have moderate to high degree of concentration. This masks significant heterogeneity across worker type, with women facing more concentrated markets than men across all ages and education levels, and less educated workers facing higher concentration than their same sex and age counterparts. Mining and quarrying, electricity, gas and steam, and water supply/sewage are the most concentrated industries.

Next, we present estimates of the labor supply parameters, labor supply elasticities and markdowns. We estimate an average labor supply elasticity across workers types and establishments of 5.790, and we find that wages are marked down roughly 17 percent below the marginal revenue product of labor. There is significant heterogeneity in the distribution of labor supply elasticities across establishments and workers, with the 10th and 90th percentiles equal to 2.800 and 8.665, respectively. Establishments that are larger in their local market tend to face a smaller labor supply elasticity and thus have more labor market power. We also find that, on average, younger workers have significantly higher elasticities than older workers, and younger women have similar or lower elasticities than younger men, while this relationship reverses for older workers. Preferences for amenities vary significantly across establishments. We find that urban areas offer more valuable amenities than rural areas, while knowledge-based and manufacturing jobs have more valuable amenities than utilities, agricultural and food service jobs. We also find that high-value amenity establishments have more workers, pay lower wages, and have lower revenue on average.

Turning to our production function estimates, we characterize the distribution of establishment and worker-type specific labor productivities and find that more educated workers have higher productivity than less educated workers, while younger workers are less productive than older workers. We find a great deal of heterogeneity in the joint distribution of establishment-specific returns to scale and total factor productivity. This joint distribution is highly skewed to the right, with a 90-10 ratio of 22.354. We also find that worker types in our setting are highly substitutable. We characterize this using the Morishima elasticity of substitution which is the appropriate elasticity concept when firms have labor market power (Morishima, 1967). Taken together, our parameter estimates indicate that establishment-specific labor demand is highly elastic, with a median labor demand elasticity of -5.317 and a range of -2.961 to -12.042 across the different labor types.

In Section 7 we perform a series of counterfactual experiments to quantitatively examine the role of labor supply and demand factors in driving wage inequality, labor market concentration, and welfare. To conduct each experiment, we begin with our estimated parameters and solve for the equilibrium distribution of wages and employment. We then remove different sources of firm and worker heterogeneity, recalculating the new counterfactual equilibrium, wage dispersion, concentration and welfare. Importantly, these counterfactual experiments take into account general equilibrium effects. Our main results highlight that all the primary channels in our model drive wage inequality. In the presence of interaction effects, the order of the decomposition matters. Some mechanisms always increase inequality (heterogeneity in worker skill) while others always decrease inequality (heterogeneity in the deterministic preference for amenities). In other cases (heterogeneity in the stochastic preference for amenities and production technology), the direction of the effect on inequality depends on which other mechanisms are active in the model. These interaction effects are primarily due to the presence of decreasing returns to scale in the production function.

Our paper relates to and builds on several strands of the literature. First, our paper builds on and contributes to the growing literature on imperfect competition in labor markets.[6] Several papers estimate firm-specific labor supply elasticities using the pass-through of firm-specific productivity or demand shocks under an assumption of monopsonistic competition with estimates typically ranging between $4-6$ (Kline et al., 2019; Dube et al., 2020; Huneeus et al., 2021; Azar et al., 2022a; Lamadon et al., 2022; Kroft et al., 2023). Berger et al. (2022a) consider an indirect inference approach that exploits changes in state-level corporate tax rates and find elasticities that range from ~ 5 (payroll-weighted average) to 9 (unweighted average). Yeh et al. (2022) estimate plant-level markdowns in the manufacturing sector in the U.S. using the "production approach" and compute a ratio of wages to MRPL of 0.65, implying that wages are marked down 35 percent below the MRPL. Staiger et al. (2010) use an exogenous change in wages at Veterans Affairs hospitals as a natural experiment and estimate a labor supply elasticity of 0.10. They also find that non-VA hospitals who were not affected by the legislated change responded by changing their own wages suggesting a role for strategic interactions. Our contribution to this literature is to formally provide an

[6]Models of imperfect competition in the labor market have recently attracted interest because of their ability to explain various labor market features, such as wage dispersion for identical workers, the correlation between firm characteristics (such as size) and wages, the lack of an impact of the minimum wage on employment, and the prevalence of gender and racial wage gaps. See Manning (2003) for an excellent overview of the literature.

identification strategy for the structural labor supply elasticity that remains valid in the presence of strategic interactions. This strategy builds on the modern approach to demand estimation in Industrial Organization (IO).

Second, we contribute to the literature on identification of production functions. Recent papers (e.g., Gandhi et al., 2020) have proposed using first-order conditions to identify the output elasticities of flexible inputs in perfectly competitive markets. We show how to identify the marginal product and output elasticities in the presence of imperfect competition and strategic interactions using data on input prices (wages) and market shares. Another contribution of our paper is to specify a rich multi-factor production function that nests several special cases considered in the literature. Our production function considers richer substitution patterns across worker types and we propose a method to identify and estimate these substitution parameters which we then characterize using the Morishima elasticity of substitution.[7]

Third, our paper closely relates to the literature on matching models. Most of the existing theoretical papers that study the existence and uniqueness of equilibrium in many-to-one matching models differ mainly in terms of whether there are transfers and whether workers are gross substitutes or complements. Kelso and Crawford (1982) consider an imperfect transferable utility (TU) model where workers are gross substitutes, while Hatfield and Milgrom (2005) extend their model to a more general framework including non-transferable utility (NTU) but do not consider complementarities. Pycia (2012) considers a many-to-one matching model with complementarities and peer effects along with ex-post Nash bargaining. It is also noteworthy that these papers and most of their extensions do not consider unobserved heterogeneity and more importantly consider a competitive market and perfect information.[8] Our contribution to this literature is to consider a wage-posting model with imperfect competition and imperfect information that allows for worker-level unobserved heterogeneity. None of the existence results in the matching literature directly apply to our context. A related paper is Azevedo (2014) who considers an imperfectly competitive, many-to-one matching market. However, Azevedo (2014) does not consider unobserved heterogeneity and mostly focuses on the case where firms compete on quantities given exogenously fixed wages. In the presence of differentiated jobs, it is more natural to assume that firms compete on wages.

[7]Lindner et al. (2022) propose a two-factor firm-level CES production function over low- and high-skilled workers. They do not identify and estimate the elasticity of substitution but rather calibrate it using external estimates.

[8]Rare exceptions which consider unobserved heterogeneity but assuming a competitive market and perfect information are Dupuy et al. (2020) and Dupuy and Galichon (2022).

Lastly, although our paper does not incorporate dynamic considerations, it relates to the search-and-matching literature which incorporate firm and worker heterogeneity. Search frictions are an important source of employer market power as emphasized by Burdett and Mortensen (1998), Postel-Vinay and Robin (2002) and Taber and Vejlin (2020).[9] Our paper is most closely related to Taber and Vejlin (2020) in terms of the broader objective of decomposing wage inequality into a skill component, a preference component, and imperfect competition. One important difference is that because our model is static, we do not consider human capital that is accumulated while working; we only allow for human capital that is exogenous and comes from investment in skills prior to working. However, while matching in most dynamic search models is one-to-one due to tractability, our static framework features many-to-one matching.[10]

2. THEORETICAL FRAMEWORK AND EMPIRICAL MODEL

Consider a static labor market with a large population of individuals divided into K finite categories/types, $k \in \{1, .., K\} \equiv \mathcal{K}$. In each category k, there are an infinite number of individuals of mass $\mathfrak{m}_k$ where $\sum_{k \in \mathcal{K}} \mathfrak{m}_k = 1$. The assumption that there are a continuum of individuals of each type is made to simplify the analysis of the existence of a stable equilibrium and also for modelling convenience.[11] In practice, the population is finite, $\mathcal{M} < \infty$. One way to rationalize this is by noting that the *proportion* of individuals in each category, $\mathfrak{m}_k \equiv \frac{m_k}{\mathcal{M}}$, in a finite population is consistent with the proportion in an infinite population. More precisely, note that $\frac{m_k}{\mathcal{M}}$ remains constant as m_k and $\mathcal{M} = \sum_{k \in \mathcal{K}} m_k$ go to infinity, where m_k denotes the number of individuals of each type k in the population and $m \equiv (m_1, ..., m_K)'$ denotes the vector of individuals in the population.

The type k itself can be thought of as being derived from a function of multiple underlying (discrete or continuous) characteristics.[12] An individual i with characteristic k is denoted by k_i. On the other side of the market, we have a finite set of firms, $\mathcal{J} \equiv \{1, .., J\}$. We do not impose the assumption that the number of firms is large and thus, we can obtain pure monopsony as a special case of the model. Firms can differentiate workers at the k level. However, within each category k, individuals can be differentiated by their unobservable

[9]Other papers in this literature include Lentz (2010); Lise et al. (2016); Hagedorn et al. (2017); Eeckhout and Kircher (2018); Lopes de Melo (2018), and Bagger and Lentz (2019).

[10]An exception is Eeckhout and Kircher (2018) who consider a frictional model with large firms.

[11]In a finite population there is almost always a profitable deviation which may complicate the analysis of the existence of a stable equilibrium.

[12]In practice, each continuous characteristic (or discrete characteristic with unbounded support) $X_d : d \in \mathcal{D}$ is transformed into a discrete random variable $\mathbf{k}_d$ with realization k_d and with finite support $\mathcal{K}_d$. Each discrete variable with finite support X_d is just relabelled $\mathbf{k}_d$. The total number of types is therefore $K = K_1 \times ... \times K_{|\mathcal{D}|}$.

(both to firms and the econometrician) characteristics and taste for different firms. Each individual i chooses to work at a firm or to be non-employed, and each firm chooses wages associated with each worker type k.

Workers: Additive Random Utility Model (ARUM). Workers have heterogeneous preferences over firms. Let the potential utility of individual i of type k if offered a wage $w_{k_ij} \equiv w_{kj} \in [0, \infty)$ to work at firm j be given by:

$$U_{ij} = \beta_{kj} \ln w_{kj} + \ln u_{kj} + \epsilon_{ij},\ j \in \{1, ..., J\}, \tag{2.1}$$

where $\ln u_{kj}$ is such that $u_{kj} \in (0, \infty)$ represents the deterministic non-pecuniary part of the worker potential utility U_{ij}, and $\beta_{kj}^{-1} \in (0, \infty)$ can be interpreted as the standard deviation of ϵ_{ij} in log-dollars. Finally, ϵ_{ij} denotes the error term (idiosyncratic payoff) which is unknown to firms. Individual i's utility of being unemployed is given by:

$$U_{i0} = \beta_{k0} \ln w_{k0} + \epsilon_{i0}, \tag{2.2}$$

where $w_{k0} \in (0, \infty)$ is the non-employment benefit which throughout this paper we consider as an observable exogenous predetermined outcome.[13] Notice that in this framework, a type k worker takes wages as given and has no market power over firms. Given the potential wage streams $\{w_{kj}\}_{0 \le j \le J}$, individual i chooses according to:

$$U_i \;=\; \max\{U_{i0}, U_{i1}, ..., U_{iJ}\} = \max_{j \in \mathcal{J} \cup \{0\}} \{v_{kj} + \epsilon_{ij}\},$$

where $v_{kj} \equiv \beta_{kj} \ln w_{kj} + \ln u_{kj}$, $v_{k0} \equiv \beta_{k0} \ln w_{k0}$. Let's denote $v_{k\cdot} \equiv (v_{k0}, v_{k1}, ..., v_{kJ})'$ and $v = (v_{1\cdot}', ..., v_{K\cdot}')'$. We can define the expected utility obtained from the choice problem, namely the *social surplus function* (McFadden, 1978, 1981), as:

$$G_{k\cdot}(v_{k\cdot}) = \mathbb{E}\Big[\max_{j \in \mathcal{J} \cup \{0\}} \{v_{kj} + \epsilon_{ij}\}\Big]. \tag{2.3}$$

In order to characterize the choice probabilities, we introduce the following regularity assumption:

Assumption 1 (Independence and absolute continuity)**.** *The joint distribution function of ϵ (i) is independent of v for all $v \in \mathcal{V} \subseteq \mathbb{R}^{K(J+1)}$, (ii) and is absolutely continuous respect to the Lebesgue measure on $\mathbb{R}^{K(J+1)}$.*

Under Assumption 1, the Williams-Daly-Zachary theorem shows that[14]

---

[13] Note that we have implicitly used a location normalization when defining potential utility. Equivalently, the utilities could have been written $U_{ij} = \ln \tilde{u}_{kj} + \beta_{kj} \ln w_{kj} + \epsilon_{ij}$, for $j \in \mathcal{J} \cup \{0\}$. However, since $\tilde{u}_{kj}$, and $\tilde{u}_{k0}$ cannot be separately identified, we directly use $u_{kj} = \frac{\tilde{u}_{kj}}{\tilde{u}_{k0}}$.

[14] See alternatively Lemma 2.1 in Shi et al. (2018a).

$$\frac{\partial G_{k\cdot}(v_{k\cdot})}{\partial v_{kj}} = \mathbb{P}\big(v_{kj} + \epsilon_{ij} \geq v_{kj'} + \epsilon_{ij'} \text{ for all } j' \in \mathcal{J} \cup \{0\} \equiv \mathcal{J}_0\big), \tag{2.4}$$

and therefore, the labor supply function is given by:

$$(\ell_{kj})^s = m_k \frac{\partial G_{k\cdot}(v_{k\cdot})}{\partial v_{kj}}, \tag{2.5}$$

where $(\ell_{kj})^s$ represents the number of type k workers that prefer firm j at the wage w_{kj}. Equation (2.5) provides a general form of labor supply that does not rely on a specific distribution of the error terms and allows for an arbitrary correlation among them. This general expression allows us to consider a general characterization of our model that does not rely on assuming that ϵ_{ij} are Type I Extreme Value (Logit).

Firms: Wage-Posting framework. Each firm j has a production function given by: $F^j(\ell_{.j})$ where $\ell_{.j} = (\ell_{1j}, ..., \ell_{kj})$. For simplicity, we ignore capital and intermediate inputs. Each firm posts a wage offer at the k level. We adopt the Bertrand-Nash assumption where each firm j chooses it's optimal wage taking other firms' wage as given. Given knowledge of the labor supply function (2.5) and exogenous output Q_j, firm j's best response consists in posting a stream of type-specific wages that minimize the cost of production. More precisely, firm j's best response is obtained as follows:[15]

$$\min_{w_{kj}} \sum_{k \in \mathcal{K}} w_{kj} \ell_{kj} \;\; s.t \; F^j(\ell_{.j}) \geq Q_j, \quad w_{kj} \geq 0$$

where

$$\ell_{kj} = m_k \frac{\partial G_{k\cdot}(v_{k\cdot})}{\partial v_{kj}}, \;\; (k, j) \in (\mathcal{K} \times \mathcal{J}). \tag{2.6}$$

Before analyzing the firm's optimal choice, we impose some regularity conditions on the production function.

Assumption 2. *(i) We assume that the minimum acceptable level of output for each firm is positive, i.e.* $Q^j > 0,\ j \in \mathcal{J}$. *(ii) We assume the firms' production functions* $F^j(.); j \in \mathcal{J}$ *to be (a) twice continuously differentiable, (b) non-constant and non-decreasing in each of their arguments, to have bounded partial derivatives, and to have zero production with zero labor inputs, i.e.* $0 \leq F^j_k(\ell_{.j}) \equiv \frac{\partial F^j(\ell_{.j})}{\partial \ell_{kj}} \leq \bar{F}' < \infty \; \forall k \in \mathcal{K}$ *and* $F^j(0) = 0$.

[15]We can equivalently consider the following minimization problem:

$$\min_{w_{kj}, (\ell_{kj})^d} \sum_{k \in \mathcal{K}} w_{kj} (\ell_{kj})^d \;\; s.t \;\; (\ell_{kj})^d = (\ell_{kj})^s, F^j(\ell_{.j}) \geq Q_j, \text{ and } w_{kj} \geq 0.$$

Let $s_{kj} \equiv \frac{\ell_{kj}}{m_k}$ denote the share of type k workers employed at firm j. Under Assumptions 1 and 2, the optimal wages (interior solutions) are given by[16]

$$w_{kj} = \lambda_j F_k^j(\ell_{\cdot j}) \frac{\mathcal{E}_{kj}}{1+\mathcal{E}_{kj}}, \quad \forall\, (k,j) \in \mathcal{K} \times \mathcal{J}, \tag{2.7}$$

where $\mathcal{E}_{kj} \equiv \frac{w_{kj}}{\ell_{kj}} \frac{\partial \ell_{kj}}{\partial w_{kj}} \equiv \frac{w_{kj}}{s_{kj}} \frac{\partial s_{kj}}{\partial w_{kj}}$ is the elasticity of labor supply at the optimal wage.[17] More generally we define the cross-wage elasticity of labor supply as: $\mathcal{E}_{kjl} \equiv \frac{w_{kl}}{s_{kj}} \frac{\partial s_{kj}}{\partial w_{kl}}$, and use the shorthand notation $\mathcal{E}_{kjj} \equiv \mathcal{E}_{kj}$. Recall that under Assumption 1, the social surplus function is convex,[18] which implies that $\mathcal{E}_{kj} \geq 0$. λ_j is the Lagrange multiplier that represents the marginal cost of production that the firm equates to marginal revenue at the optimal choice of Q_j; hereafter we assume that λ_j is bounded, i.e. $0 < \lambda_j < \bar{\lambda} < \infty$. Notice that Assumption 2 (ii-b) ensures that $F_k^j(\ell_{\cdot j}) \geq 0$, so the optimal wage is always non-negative. Let $\mathcal{C}^j \subseteq \mathcal{K}$ denote the set of worker types for whom firm j offers a strictly positive wage, $w_{kj} > 0$ which according to our ARUM specification and Assumption 1 is equivalent to $s_{kj} > 0$. Assumption 2 (ii-b) and the optimality conditions ensure that for all firms j in this market $\mathcal{C}^j \neq \{\emptyset\}$, where $\mathcal{C}^j \equiv \{k \in \mathcal{K} : s_{kj} > 0\} = \{k \in \mathcal{K} : w_{kj} > 0\}$.

Furthermore, we can write the labor supply elasticity in terms of the social surplus function as follows:

$$\mathcal{E}_{kj} = \beta_{kj} \frac{\frac{\partial^2 G_{k\cdot}(v_{k\cdot})}{\partial^2 v_{kj}}}{\frac{\partial G_{k\cdot}(v_{k\cdot})}{\partial v_{kj}}}.$$

Therefore, each firm plays it best response strategy taking other firms' wage as given whenever their posted wage stream is given as follows:

$$w_{kj} = \lambda_j \beta_{kj} F_k^j(\ell_{\cdot j}) \frac{\frac{\partial^2 G_{k\cdot}(v_{k\cdot})}{\partial^2 v_{kj}}}{\frac{\partial G_{k\cdot}(v_{k\cdot})}{\partial v_{kj}} + \beta_{kj} \frac{\partial^2 G_{k\cdot}(v_{k\cdot})}{\partial^2 v_{kj}}} \forall\, (k,j) \in \mathcal{C}^j \times \mathcal{J}. \tag{2.8}$$

So far we have described the behaviour of each side of the market. Now, we define an equilibrium for this many-to-one employee-employer matching model. Let $\mathbb{R}_{\geq 0}$ denote $\{x \in \mathbb{R} : x \geq 0\}$ and $\mathbb{R}_{>0} \equiv \{x \in \mathbb{R} : x > 0\}$.

Definition 1. *Consider workers that have preferences which are of the ARUM form, i.e. eq. (2.1) and firms that have production functions which satisfy Assumption 2. An equilibrium outcome* (s, w) *consists of a distributional worker-firm matching function and an equilibrium wage equation such that* $w \equiv (w_{10}, ..., w_{KJ}) \in (\mathbb{R}_{\geq 0})^{K(J+1)}$ *and* $s \equiv (s_{10}, ..., s_{KJ}) \in [0,1]^{K(J+1)}$ *are optimal for workers and firms (workers maximize their utilities, firms set*

[16]The details of the derivation are relegated to Appendix A.1

[17]By convention and to ease the notation, we consider that $\mathcal{E}_{kj} = 0$ when $s_{kj} = 0$.

[18]See McFadden in Manski et al. (1981), or Shi et al. (2018b), Lemma 2.1.

their optimal wages in a Bertrand oligopsony model), and the following population constraint holds

$$\sum_{j\in\mathcal{J}} s_{kj} + s_{k0} = 1, \quad k \in \mathcal{K}. \tag{2.9}$$

Under Assumptions 1, and 2, the equilibrium outcome is equivalent to satisfying equations (2.6), (2.8) and (2.9). In the next section, we formally derive the conditions under which such an equilibrium exists and is unique.

3. EXISTENCE AND UNIQUENESS OF EQUILIBRIUM

In this section, we establish existence and uniqueness of equilibrium for a wide class of many-to-one matching models in presence of unobserved heterogeneity on the workers side and imperfect competition.

Theorem 1. *[Existence] Under Assumptions 1 and 2, an equilibrium exists.*

The proof presented formally in Appendix B.1 mainly relies on Brouwer's fixed-point theorem. In a many-to-one matching model with a finite number of firms and unrestricted strategic interactions, a shock to one firm's productivity could affect the equilibrium employment and wages of other firms; therefore, the existence of multiple equilibria should not be surprising in such an environment.[19] However, we now characterize a set of shape restrictions on the firms' production functions and the labor supply elasticities that ensures the existence of a unique equilibrium. First, we define the k type "cross-wage super-elasticities" of labor supply as: $\zeta_{kjl} \equiv \frac{w_{kl}}{\mathcal{E}_{kj}}\frac{\partial \mathcal{E}_{kj}}{\partial w_{kl}}$. ζ_{kjl} is the elasticity of the labor supply elasticity of type k worker at firm j with respect to the type k wage at firm l, w_{kl}. In the absence of strategic interactions, $\zeta_{kjl} = 0$ for $j \neq l$. Also, notice that $\zeta_{kjj} \equiv \zeta_{kj}$ is the so-called "super-elasticity" discussed in Klenow and Willis (2016), Nakamura and Zerom (2010), and Edmond et al. (2023).

Assumption 3 (Shape Restrictions)**.** *(i) [cross-wage super-elasticity] Assume that we have a social surplus function such that whenever all others entries $w_{kl'}$ for $l' \neq l$, remain constant we have for all $k \in \mathcal{K}$*

$$\zeta_{kjl} \begin{cases} \leq 0, & \text{if } l = j \\ \geq 0, & \text{if } l \in \mathcal{J}_0 \setminus \{j\} \end{cases}$$

[19] Card et al. (2018) also discuss the complications that arise in the presence of multiple equilibria in a framework with a finite number of firms.

(ii) [Production function] The production function takes the following functional form:

$$F^j(\ell_{\cdot j}) = \sum_{k \in \mathcal{K}} h_k(\ell_{kj}),$$

where h is a $C^2(\mathbb{R})$ function such that $h'_k(x) \geq 0$ and $h''_k(x) \leq 0$.

Assumption 3 (i) imposes sign restrictions on the cross-wage super-elasticities. The sign restriction imposes that when firm j increases the wage of type k workers, the labor supply elasticity decreases; on the other hand, it increases when another firm l increases the type k wage. This sign restriction is satisfied for a wide class of error distributions, including the Nested Logit that we analyze in detail below. The restriction imposed on the production function—Assumption 3 (ii)—allows for decreasing or constant returns to scale, and for a non-constant marginal rate of substitution. A special case of Assumption 3 (ii) is: $F^j(\ell_{\cdot j}) = \theta_j\Big(\sum_{k\in\mathcal{K}} \gamma_{kj}\ell_{kj}^{\rho_{kj}}\Big)$ where $\theta_j > 0$ is total factor productivity, and $\gamma_{kj} > 0$ are such that $\sum_{k\in\mathcal{K}} \gamma_{kj} = 1$. $\rho_{kj} \in (0, 1]$ for $k \in \mathcal{K}$ parametrize the marginal rate of substitution between different type of workers at firm j.[20] Some papers, e.g., Postel-Vinay and Robin (2002), consider skill as one dimensional and not varying across firms, i.e. $\gamma_{kj} = \gamma_k$. The functional form we entertain above does not impose these restrictions and instead follows Roy (1951) and more recently Taber and Vejlin (2020) by allowing for worker-employer match-specific productivity, whereby a specific type of worker may be more productive in some firms compared to other firms.

Theorem 2. *[Existence and Uniqueness] Under Assumptions 1, 2, and 3 an equilibrium exists and it is unique.*

The proof relies on the observation that the mapping induced by eq (2.7) is globally invertible, since its Jacobian matrix is positive diagonally dominant. For the sake of simplicity, the detailed proof is relegated to Appendix B.2. In the proof, we also discuss the case where the production function is not necessarily additive separable, i.e. Assumption 3 (ii) does not hold. In such a context, we show that the equilibrium can also be unique under an additional sign restriction on a component involving the production function partial mixed-derivatives and the cross-wage labor supply elasticity, i.e. $F^j_{kl}(\ell_{\cdot j}) \equiv \frac{\partial^2 F^j(\ell_{\cdot j})}{\partial \ell_{kj}\partial \ell_{kl}}$ and $\mathcal{E}_{kjl}$. This restriction could be tested if the primitive parameters of this model are known or identified. The uniqueness result is important when performing counterfactual analyses.

Special case: Nested Logit Economy. To allow unobserved workers preferences ϵ_{ij} to be correlated for certain classes of firms, we partition the J firms into G nests, the g^{th} nest

[20]Another special case is $F^j(\ell_{\cdot j}) = \theta_j + \ln\Big(\Pi_{k\in\mathcal{K}}\ell_{kj}^{\gamma_{kj}}\Big)$ which is the log-linearization of the well known Cobb-Douglas production function.

containing N_g firms. In our empirical application, we consider a nest to be the local labor market and define it as all firms belonging to the same community zone and industry. We assume the ϵ_{ij} to be correlated within nests, i.e. $1/\sigma_{kg} = \sqrt{1 - corr(\epsilon_{ij}, \epsilon_{il})}$ for $j \neq l$ where for $(j,l) \in N_g$, and with $\sigma_{kg} \in [1, \infty)$. Despite the nesting structure, we allow each firm to compete with every other firm in the economy, regardless of whether firms belong to the same nest or not. In this Nested Logit Economy, the social surplus function is given by

$$G_{k\cdot}(v_{k\cdot}) = \ln\Big\{ \overbrace{e^{v_{k0}} + \sum_{g=1}^{G} \Big(\underbrace{\sum_{j \in N_g} e^{v_{kj}\sigma_{kg}}}_{\mathcal{I}_{k,g}(v_{k\cdot})} \Big)^{1/\sigma_{kg}}}^{\mathcal{I}_{k,M}(v_{k\cdot})} \Big\},$$

where $\mathcal{I}_{k,g}(v_{k\cdot})$ and $\mathcal{I}_{k,M}(v_{k\cdot})$ denote, respectively, the aggregate weighted wage index at the nest g level, and at the whole market level. Additionally, the market shares have the following weakly separable functional form: $s_{kj}(w_{k\cdot}) = f\big(w_{kj}, \mathcal{I}_{k,g}(v_{k\cdot}), \mathcal{I}_{k,M}(v_{k\cdot})\big)$. In this case, the labor supply elasticities are given by:

$$\begin{aligned} \mathcal{E}_{kj} = \frac{w_{kj}}{s_{kj}} \Bigg[f_1\big(w_{kj}, \mathcal{I}_{k,g}(v_{k\cdot}), \mathcal{I}_{k,M}(v_{k\cdot})\big) \quad &+ \quad \frac{\partial \mathcal{I}_{k,g}(v_{k\cdot})}{\partial w_{kj}} f_2\big(w_{kj}, \mathcal{I}_{k,g}(v_{k\cdot}), \mathcal{I}_{k,M}(v_{k\cdot})\big) \\ &+ \quad \frac{\partial \mathcal{I}_{k,M}(v_{k\cdot})}{\partial w_{kj}} f_3\big(w_{kj}, \mathcal{I}_{k,g}(v_{k\cdot}), \mathcal{I}_{k,M}(v_{k\cdot})\big) \Bigg] \end{aligned}$$

where $f_k(x_1, x_2, x_3) = \frac{\partial f(x_1,x_2,x_3)}{\partial x_k}$ for $k \in \{1, 2, 3\}$. The last equality shows that a change in w_{kj} has a direct effect on the share s_{kj} captured by $f_1(.)$ and two indirect effects mediated by the impact of the change of w_{kj} on the local and the total market indexes $\mathcal{I}_{k,g}(v_{k\cdot})$, and $\mathcal{I}_{k,M}(v_{k\cdot})$, respectively. Card et al. (2018) and Lamadon et al. (2022) consider a special case of imperfect competition which implies that the two latter effects are null, i.e. $\frac{\partial \mathcal{I}_{k,g}(v_{k\cdot})}{\partial w_{kj}} f_2(.) + \frac{\partial \mathcal{I}_{k,M}(v_{k\cdot})}{\partial w_{kj}} f_3(.) = 0$. Such an assumption can considerably limit the effect of the market power for some firms and impose important restrictions on the nature of strategic interactions. For instance, these frameworks assume away the possibility that some firms are dominant in a certain local market g, in such a way that they may hire a non-negligible share of some types of workers in their local market. Under this assumption, productivity or amenities shocks in firm j that affect w_{kj} do not have any spillover effects onto the equilibrium wage in a different firm j', $w_{kj'}$. Berger et al. (2022a), on the other hand, impose the weaker condition that $\frac{\partial \mathcal{I}_{k,M}(v_{k\cdot})}{\partial w_{kj}} f_3(.) = 0$; in other words, they allow some firms to be dominant in their local market but no firm has enough power to hire a significant share of some type of workers at the aggregate market level.[21] In this paper, we do not impose any

[21]In their context, this restriction arises they consider a model with an infinite number of local markets.

of these restrictions. In particular, the labor supply elasticity in the Nested Logit economy takes the following form:

$$\mathcal{E}_{kj} \quad = \quad \beta_{kj}[\sigma_{kg} + (1-\sigma_{kg})s_{kj|g} - s_{kj}] \quad \text{for } j \in N_g \tag{3.1}$$

with $s_{kj} \equiv e^{v_{kj}\sigma_{kg}}\mathcal{I}_{k,g}(v_{k\cdot})^{1/\sigma_{kg}-1}\mathcal{I}_{k,M}(v_{k\cdot})^{-1}$, $s_{kj} \equiv e^{v_{kj}\sigma_{kg}}\mathcal{I}_{k,g}(v_{k\cdot})^{1/\sigma_{kg}-1}\mathcal{I}_{k,M}(v_{k\cdot})^{-1}$, $s_{kg} = \sum_{j\in N_g} s_{kj} = \mathcal{I}_{k,g}(v_{k\cdot})^{1/\sigma_{kg}}\mathcal{I}_{k,M}(v_{k\cdot})^{-1}$, and $s_{kj|g} = \frac{s_{kj}}{s_{kg}} = e^{v_{kj}\sigma_{kg}}\mathcal{I}_{k,g}(v_{k\cdot})^{-1}$ where $s_{kj|g}$ denotes the share of workers of type k working in the firm j as a fraction of the total nest share. Note that the atomistic firm assumption considered in Card et al. (2018) and Lamadon et al. (2022) implies that $(1-\sigma_{kg})s_{kj|g} - s_{kj} = 0$ for all $(k,j) \in \mathcal{K} \times \mathcal{J}$, and $g \in \{1, ..., G\}$, meaning that even at the local market $s_{kj|g} = s_{kj} = 0$, provided that $\sigma_{kg} > 1$. Therefore, if we observe in the data that some firms have a significant share of type k workers in their local market, i.e. $s_{kj|g} > \underline{s}$ for $\underline{s} > 0$ we can reject the atomistic firm assumption. Another important remark is that we always have $[(1-\sigma_{kg})s_{kj|g} - s_{kj}] \leq 0$, which implies that the atomistic firm assumption leads to an overestimation of firms' labor supply elasticities and then the markdowns. Berger et al. (2022a)'s restriction imposes that $s_{kj} = 0$ for all (k,j), but allows $(1-\sigma_{kg})s_{kj|g} \neq 0$ for some (k,j). Therefore, they also tend to overestimate the true markdowns but with a lower bias than the one estimated under the atomistic firm assumption.[22]

The cross-wage super-elasticities in the Nested Logit model take the following form:

$$\zeta_{kjl} = \beta_{kj}\left[(1-\sigma_{kg})s_{kj|g}\frac{\mathcal{E}_{kjl|g}}{\mathcal{E}_{kj}} - s_{kj}\frac{\mathcal{E}_{kjl}}{\mathcal{E}_{kj}}\right] \tag{3.2}$$

where $\mathcal{E}_{kjl|g}$ denotes the within-nest cross-wage elasticities. The super-elasticity simplifies to:[23]

$$\zeta_{kj} = \beta_{kj}\left[\beta_{kj}(1-\sigma_{kg})s_{kj|g}(1-s_{kj|g})/\mathcal{E}_{kj} - s_{kj}\right]. \tag{3.3}$$

Note that both the atomistic firm assumption and Berger et al. (2022a)'s restriction lead to an overestimation of the super-elasticity.

A direct application of Theorem 2 leads to the following result:

Corollary 1. *Whenever Assumptions 1, 2, and 3 (ii) hold and workers idiosyncratic utility shocks have a Nested Logit structure, an equilibrium exists and it is unique.*

The proof is immediate by showing that the sign restriction in Assumption 3 (i) holds in the Nested Logit Economy.

[22] When firms compete according to Bertrand, the labor supply elasticity in Berger et al. (2022a) is given by: $\mathcal{E}_{kj} = [\theta s_{kj|g} + \eta(1-s_{kj|g})]$ which is a special case of our elasticity when $\theta = \beta_{kj}$, $\eta = \beta_{kj}\sigma_{kg}$ and $s_{kj} = 0$.

[23] We could write also the elasticity as a function of the super-elasticity as in Edmond et al. (2023), i.e. $\mathcal{E}_{kj} = \frac{\zeta_{kj} + \beta_{kj}s_{kj}}{\beta_{kj}^2(1-\sigma_{kg})s_{kj|g}(1-s_{kj|g})}$.

3.1. **Finding the Equilibrium: An iterative method.** Solving the model equilibrium is very important to perform counterfactual analyses. Here we establish conditions under which there exist globally convergent methods for recovering the unique equilibrium outcome (s, w). Let's define

$$\delta_{kj}(w) \equiv w_{kj} - \lambda_j F_k^j\left(\ell_{\cdot j}(w)\right) \frac{\mathcal{E}_{kj}(w)}{1+\mathcal{E}_{kj}(w)}, \quad \forall (k,j) \in \mathcal{K} \times \mathcal{J}. \tag{3.4}$$

$\delta(w) = (\delta_{11}(w), ..., \delta_{KJ}(w)) : \mathbb{T}_\epsilon \subseteq \mathbb{R}^{KJ} \longrightarrow \mathbb{R}^{KJ}$, where $\mathbb{T}_\epsilon$ is a closed and bounded rectangular region.[24]

Algorithm 1 (Underrelaxed Gauss-Seidel Iteration)**.** *For $\xi \in (0,1]$:*

(1) Solve $\delta_{kj}(w_{11}^{t+1}, ..., w_{1J}^{t+1}, ..., w_{k,j-1}^{t+1}, w_{kj}, w_{k,j+1}^{t}, ..., w_{KJ}^{t}) = 0$ for w_{kj} holding all other components fixed.

(2) Set $w_{kj}^{t+1} = (1-\xi)w_{kj}^t + \xi w_{kj}$ and this for $kj = 11, ..., KJ$ and $t = 0, 1, ...$

Algorithm 2 (Underrelaxed Jacobi Iteration)**.** *For $\xi \in (0,1]$:*

(1) Solve $\delta_{kj}(w_{11}^{t}, ..., w_{1J}^{t}, ..., w_{k,j-1}^{t}, w_{kj}, w_{k,j+1}^{t}, ..., w_{KJ}^{t}) = 0$ for w_{kj} holding all other components fixed.

(2) Set $w_{kj}^{t+1} = (1-\xi)w_{kj}^t + \xi w_{kj}$ and this for $kj = 11, ..., KJ$ and $t = 0, 1, ...$

Proposition 1 (Convergence of the nonlinear Gauss-Seidel and Jacobi iteration)**.** *Consider that Assumptions 1, 2, and 3 hold. For $\xi \in (0,1]$ and any initial value $w^0 \in \mathbb{T}_\epsilon$ the nonlinear Gauss-Seidel or Jacobi iteration described in Algorithms 1 and 2 converge to the unique equilibrium wage w^{eq}. Then the equilibrium outcome is given by (w^{eq}, s^{eq}) with $s_{kj}(w^{eq}) = \frac{\partial G_{k\cdot}(v_{k\cdot})}{\partial v_{kj}}|_{v_{kj}=v_{kj}^{eq}}$ where $v_{kj}^{eq} \equiv \beta_{kj} \ln w_{kj}^{eq} + \ln u_{kj}$.*

The proof is relegated to Appendix B.3.[25]

3.2. **Comparative Statics.** In the previous subsection we proposed an efficient computational method to allow the researcher to undertake counterfactual analysis using our general framework. However, deriving analytic results on the effect of exogenous changes of some model parameters such as changes in the non-employment benefit, TFP or amenities on equilibrium wages is useful because these results provide insight into the economic structure of the model. In a nonlinear system of equations, it is generally challenging to derive comparative statics. It typically involves an application of the Implicit Function Theorem, which requires deriving a closed-form of the inverse of the Jacobian matrix associated with

[24] Please refer to the proof of Theorem 1 in Appendix B.1 for the complete definition of $\mathbb{T}_\epsilon$.

[25] A key advantage of this result is that the Gauss-Seidel and Jacobi algorithms are easy to implement and can attain fairly quick convergence, even with very large systems of equations.

the mapping defined in eq (3.4). In the presence of strategic interactions, obtaining a closed-form of this inverse matrix is quite challenging, especially when the number of firm is large. However, we show that this Jacobian matrix has some special features that allow us to find informative bounds.

Recall that

$$w_{kj} = \lambda_j \underbrace{F_k^j(\ell_{\cdot j})}_{\mathrm{mpl}_{kj}} \underbrace{\frac{\mathcal{E}_{kj}}{1+\mathcal{E}_{kj}}}_{\mathrm{md}_{kj}}, \ \ \forall\,(k,j) \in \mathcal{K} \times \mathcal{J}.$$

where mpl_{kj} and md_{kj} denote, respectively, the marginal productivity of labor and the markdown of firm j for a type k worker. The elasticity of mpl_{kj} and md_{kj} respect to w_{kl} are given by:

$$\frac{\partial \ln \mathrm{mpl}_{kj}}{\partial \ln w_{kl}} = \underbrace{\frac{w_{kl}}{\ell_{kj}} \frac{\partial \ell_{kj}(w_{k\cdot})}{\partial w_{kl}}}_{\mathcal{E}_{kjl}} \underbrace{\left(\frac{F_{kk}^j}{F_k^j} \ell_{kj} \right)}_{1/\eta_{kj}},$$

$$\frac{\partial \ln \mathrm{md}_{kj}}{\partial \ln w_{kl}} = \underbrace{\frac{1}{(1+\mathcal{E}_{kj}(w_{k\cdot}))}}_{1-\mathrm{md}_{kj}} \underbrace{\frac{w_{kl}}{\mathcal{E}_{kj}(w_{k\cdot})} \frac{\partial \mathcal{E}_{kj}(w_{k\cdot})}{\partial w_{kl}}}_{\zeta_{kjl}}.$$

Recall that λ_j is the marginal cost of output. At the profit maximizing level of output, λ_j is equal to marginal revenue product. For the sake of simplicity, in the following proposition we assume that all firms j are price takers on the output market, $\lambda_j = P_j$ where P_j is the exogenous price.[26] Under this assumption, we can define the "labor demand elasticity" as the elasticity of the inverse marginal revenue product of labor curve:[27]

$$\eta_{kj} \equiv \frac{F_k^j}{\ell_{kj} F_{kk}^j}. \tag{3.5}$$

The cross-wage elasticities $\mathcal{E}_{kjl}$, the cross-wage super-elasticities ζ_{kjl}, the markdowns md_{kj}, and the labor demand elasticity η_{kj}, are the key statistics that drive our comparative statics results. They are the key channels by which an exogenous shock at firm l affects firm j's equilibrium wage. Recall that under the atomistic firms assumption imposed in Card et al. (2018) and Lamadon et al. (2022), $\mathcal{E}_{kjl} = 0$ for all $l \neq j$ and $\zeta_{kjl} = 0$ for all $l, j \in \mathcal{J}$. The equilibrium restriction entertained in Berger et al. (2022a)'s relaxes the latter restrictions but still imposes that $\mathcal{E}_{kjl} = \zeta_{kjl} = 0$ for all firms l and j belonging to

[26] It is worth noting this restriction is not critical for deriving $\frac{\theta_l}{w_{kj}} \frac{\partial w_{kj}}{\partial \theta_l}$.

[27] See Weyl and Fabinger (2013) who define a similar object when analyzing the output market, although note that in their setting, the relevant object is the output "supply elasticity".

different local markets or groups. We do not impose such restrictions and thus provide a much general set of comparative statics. Before presenting our main results, let us consider the following shorthand notation:

$$\psi_{k,jl} \equiv \frac{\mathcal{E}_{kjl}}{\eta_{kj}} + (1 - \mathrm{md}_{kj})\zeta_{kjl}.$$

Analogous, to $\psi_{k,jl}$ we also define $\phi_{k,jl} = \frac{\partial \ln \mathrm{mpl}_{kj}}{\partial \ln u_{kl}} + \frac{\partial \ln \mathrm{md}_{kj}}{\partial \ln u_{kl}}$ which are the elasticities respect to u_{kl}. In the next result, we derive closed-form comparative statics for the case of duopsony and lower bounds for the general oligopsony case.

Proposition 2 (Comparative Statics). *Consider that Assumptions 1, 2, and 3 hold. Let (s, w) denotes the unique equilibrium outcome of our many-to-one matching model. In a neighbourhood of the equilibrium (s, w) the following (general equilibrium) comparative statics hold:*

(i) **Duopsony**: *$\mathcal{J} = \{j, l\}$. For any $k \in \mathcal{C}^j \cap \mathcal{C}^l$ we have*

(a)

$$\frac{w_{k0}}{w_{kj}} \frac{\partial w_{kj}}{\partial w_{k0}} = \frac{(1 - \psi_{k,ll})\psi_{k,j0} + \psi_{k,jl}\psi_{k,l0}}{(1 - \psi_{k,jj})(1 - \psi_{k,ll}) - \psi_{k,jl}\psi_{k,lj}} \geq 0.$$

(b) If the firms' production functions have a multiplicative structure $F^l(.) = \check{\theta}_l \check{F}^l(.)$ where $\frac{\partial \check{F}^l(.)}{\partial \check{\theta}_l} = 0$ then for any $k \in \mathcal{C}^j \cap \mathcal{C}^l$ we have

$$\frac{\check{\theta}_l}{w_{kj}} \frac{\partial w_{kj}}{\partial \check{\theta}_l} = \frac{\psi_{k,jl}}{(1 - \psi_{k,jj})(1 - \psi_{k,ll}) - \psi_{k,jl}\psi_{k,lj}} \geq 0,$$

$$\frac{\check{\theta}_l}{w_{kl}} \frac{\partial w_{kl}}{\partial \check{\theta}_l} = \frac{(1 - \psi_{k,jj})}{(1 - \psi_{k,jj})(1 - \psi_{k,ll}) - \psi_{k,jl}\psi_{k,lj}} > 0.$$

(c)

$$\frac{u_{kl}}{w_{kj}} \frac{\partial w_{kj}}{\partial u_{kl}} = \frac{(1 - \psi_{k,ll})\phi_{k,jl} + \psi_{k,jl}\phi_{k,ll}}{(1 - \psi_{k,jj})(1 - \psi_{k,ll}) - \psi_{k,jl}\psi_{k,lj}} \gtreqless 0,$$

$$\frac{u_{kl}}{w_{kl}} \frac{\partial w_{kl}}{\partial u_{kl}} = \frac{(1 - \psi_{k,jj})\phi_{k,ll} + \psi_{k,lj}\phi_{k,jl}}{(1 - \psi_{k,jj})(1 - \psi_{k,ll}) - \psi_{k,jl}\psi_{k,lj}} \gtreqless 0.$$

(ii) **Oligopsony**: *$J \geq 2$. For any $k \in \mathcal{C}^j \cap \mathcal{C}^l$, and $l, j \in \mathcal{J}$, we have*

(a) For any $k \in \mathcal{C}^j$ we have:

$$\frac{w_{k0}}{w_{kj}} \frac{\partial w_{kj}}{\partial w_{k0}} \geq \frac{\mathcal{E}_{kj0}/\eta_{kj} + (1 - md_{kj})\zeta_{kj0}}{1 - \mathcal{E}_{kj}/\eta_{kj} - (1 - md_{kj})\zeta_{kj}} \geq 0.$$

(b) If the firms' production functions have a multiplicative structure $F^l(.) = \check{\theta}_l \check{F}^l(.)$ *where* $\frac{\partial \check{F}^l(.)}{\partial \check{\theta}_l} = 0$ *then for any* $k \in \mathcal{C}^j \cap \mathcal{C}^l$ *we have:*

$$\frac{\check{\theta}_l}{w_{kj}} \frac{\partial w_{kj}}{\partial \check{\theta}_l} \begin{cases} \geq \frac{\mathcal{E}_{kjl}/\eta_{kj} + (1-md_{kj})\zeta_{kjl}}{\left(1-\mathcal{E}_{kj}/\eta_{kj}-(1-md_{kj})\zeta_{kj}\right)(1-\mathcal{E}_{kl}/\eta_{kl}-(1-md_{kl})\zeta_{kl})} \geq 0 \text{ if } j \neq l, \\ \geq \frac{1}{(1-\mathcal{E}_{kl}/\eta_{kl}-(1-md_{kl})\zeta_{kl})} > 0, \text{ if } j = l. \end{cases}$$

where $\psi_{k,jl}, \phi_{k,jl} \geq 0$ *for* $l \neq j$*, and* $\psi_{k,ll}, \phi_{k,ll} \leq 0$.

Before discussing the results, it is worth noting that all the lower bounds are sharp, in the sense that there exists a data generating process under which these inequalities hold exactly as equality. As we discuss below, the inequalities present in Proposition 2 (ii) hold as equalities when strategic interactions are assumed away.[28]

Non-employment benefit shocks. Proposition 2:(i)/(ii)-a shows the resulting effect of an exogenous increase of non-employment benefits on the equilibrium wages. The equation in (i)-a shows explicitly the different channels by which an exogenous shock to non-employment benefits affects the equilibrium wages in the duopsony case: An increase of w_{k0} has a direct effect on mpl_{kj} and md_{kj}, and firm j adjusts w_{kj} in response. An indirect effect is transmitted through firm l: the increase of w_{k0} also has a direct effect on mpl_{kl} and md_{kl}, and firm l adjusts w_{kl}. This change in w_{kl} affects firm j through $\psi_{k,jl}$, firm j then responds by changing w_{kj}, and this in turn generates a response of firm l through $\psi_{k,lj}$. This succession of responses converges and leads to a final increase of w_{kj}. In sum, the strategic responses are mediated by $\psi_{k,jl}$ and $\psi_{k,lj}$ in this duopsony context. In the more general case with $J \geq 2$, the strategic interactions are captured by $\psi_{k,jr}$ and $\psi_{k,rj}$ for all $r \in \mathcal{J} \setminus \{j\}$. Proposition 2:(ii)-a shows that those indirect effects due to strategic interactions only amplify the magnitudes of the effect of an exogenous increase of non-employment benefits on the equilibrium wages. Indeed, the lower bound derived in (ii)-a is achieved when there are no strategic interactions, i.e., $\psi_{k,jr} = \psi_{k,rj} = 0$ for all $r \in \mathcal{J} \setminus \{j\}$, which happens for example under the "atomistic" firms assumption imposed in Card et al. (2018) and Lamadon et al. (2022) or in the Berger et al. (2022a) framework where each local market contains only one firm.

TFP shocks. In Proposition 2: (i)/(ii)-b we consider the effect of a positive increase of the total factor of production (TFP) of firm l on the equilibrium wages in the economy. We assume that firm l's production function takes the form $F^l(.) = \check{\theta}_l \check{F}^l(.)$ where $\frac{\partial \check{F}^l(.)}{\partial \check{\theta}_l} = 0$, and that $\check{F}^l(.)$ respects Assumption 3 (iii). Equations in (i)-b show again two key channels

[28] A formal proof of this statement is derived in the proof of Proposition 2 relegated in Appendix B.4.

by which a productivity shock in firm l affects the equilibrium wage in a duopsony market. The increase in $\check{\theta}_l$ has a direct effect on mpl_{kl} and firm l adjusts w_{kl} through $\psi_{k,ll}$. This then affects mpl_{kj} and md_{kj} through $\psi_{k,jl}$, and firm j then responds to this change through $\psi_{k,jj}$. This succession of responses converges to a final increase of w_{kj} and w_{kl}. Notice that unlike the unemployment benefit comparative static, in this case, the only way a $\check{\theta}_l$ shock is initially transmitted to firm j is through $\psi_{k,jl}$. In the more general case ($J \geq 2$), (ii)-b shows again that strategic interactions amplify the magnitude of the effect of an exogenous productivity shock on equilibrium wages. Indeed, the lower bound for $\frac{\check{\theta}_l}{w_{kl}} \frac{\partial w_{kl}}{\partial \check{\theta}_l}$ is attained when all the strategic interactions are assumed away. This situation encompasses two interesting special cases: (i) A framework where there is large number of local markets with a single dominant firm in each local market. These dominant firms do not internalize the impact of their wage setting on the whole market wage index, but internalize it at the local market level. In this case, the labor supply elasticity of each firm j is variable and depends on its own market shares. This is a special case of the framework in Berger et al. (2022a). (ii) The monopsonistic competition framework considered in Card et al. (2018) and Lamadon et al. (2022), where the labor supply elasticities of all firms are constant.

To clarify how our comparative statics results generalize the special cases analyzed in the literature, we consider the Nested Logit Economy. In this case, the lower bound of Proposition 2: (ii)-b simplifies to:

$$\left\{ \underbrace{\underbrace{1 - \frac{\beta_{kj}\sigma_{kg}}{\eta_{kj}}}_{LMS} - \beta_{kj}(1-\sigma_{kg}) s_{kj|g} \left[\frac{1}{\eta_{kj}} + \beta_{kj}(1 - s_{kj|g}) \frac{(1-\text{md}_{kj})^2}{\text{md}_{kj}} \right]}_{BHM} + \beta_{kj} s_{kj} \left[\frac{1}{\eta_{kj}} + (1 - \text{md}_{kj}) \right] \right\}^{-1} \tag{3.6}$$

LMS denotes the passthrough formula obtained in Lamadon et al. (2022) where firms are atomistic, i.e. $s_{kj|g} = s_{kj} \approx 0$. BHM represents the passthrough formula in the Berger et al. (2022a) framework where strategic interactions channels are shut down, i.e. only one dominant firm per local market.[29] Here, our lower bound provides the general formula for the passthrough when all cross-wage elasticities and cross-wage super-elasticities are assumed to be zero, i.e. $\mathcal{E}_{kjl} = \zeta_{kjl} = 0$ for $l \neq j$, i.e., shutting down all strategic interaction channels. No specific restrictions are imposed on $\mathcal{E}_{kj}$ and ζ_{kj}.

Amenities shocks. In Proposition 2: (i)-b we analyze the effect of a positive increase of type k worker preference for firm l amenities on the equilibrium wages. The analysis of the duopsony shows that in the case of an amenities shock, the indirect effect due to strategic interactions works against the direct effect and does not allows us to determine the sign

[29]Notice that in the Berger et al. (2022a) case, the markdown is restricted to the case where $s_{kj} = 0$.

of the equilibrium effect. The increase of u_{kl} directly affects mpl_{kl} and md_{kl} through $\phi_{k,ll}$, the firm then adjusts the wage w_{kl}, and this change affects firm j through $\psi_{k,jl}$. However, at the same time, the change in u_{kl} directly affects mpl_{kj} and md_{kj} through $\phi_{k,jl}$, firm j responds, and after a set of iterative responses we have the final effect. As can be seen when the strategic interaction terms are 0, i.e. $\psi_{k,jl} = \psi_{k,lj} = 0$, we have $\frac{u_{kl}}{w_{kl}} \frac{\partial w_{kl}}{\partial u_{kl}} < 0$. But when $\psi_{k,jl}$, $\psi_{k,lj}$ are not null, the resulting aggregate effect could be positive.

4. Social Welfare, Generalized Entropy and Market Concentration.

In this section, we define social welfare and establish a link to market concentration. We assume that total firm profits in the economy are redistributed in the form of payments to a group $\mathcal{R} \subseteq \mathcal{K} \times \mathcal{J}_0$ of agents, in proportion to their equilibrium wages (non-employment benefit for the non-employed). More precisely, we have

$$\sum_{j=1}^{J} \underbrace{\left(\lambda_j F^j(\ell_{\cdot j}) - \sum_{k=1}^{K} w_{kj} \ell_{kj} \right)}_{\pi_j} = \sum_{(k,j) \in \mathcal{R}} \phi(s, w; \lambda, \mathcal{R}) w_{kj} \ell_{kj}, \tag{4.1}$$

where $\lambda = (\lambda_1, ..., \lambda_J)'$. Let's collect all of the primitives parameters of the model into a vector Ξ. The social welfare function for the many-to-one matching model is defined as an adjusted version of the social surplus function (utilitarian social welfare function):[30]

$$\mathcal{W}(\Xi, \lambda, \mathcal{R}) = \sum_{k=1}^{K} m_k G_{k\cdot}(\widetilde{v}_{k\cdot}) \tag{4.2}$$

where[31]

$$\widetilde{v}_{kj} \equiv \begin{cases} \beta_{kj} \ln\left\{ w_{kj}(1 + \phi(s, w; \lambda, \mathcal{R})) \right\} + \ln u_{kj} = v_{kj} + \beta_{kj} \ln(1 + \phi(s, w; \lambda, \mathcal{R})), \text{ if } (k,j) \in \mathcal{R} \\ v_{kj}, \text{ if } (k,j) \notin \mathcal{R}. \end{cases}$$

In this representation, all agents that are not included in $\mathcal{R}$ are excluded from the profit sharing. Let $G^*_{k\cdot}(s_{k\cdot})$ denote the *convex conjugate* or *Legendre-Fenchel transform* of $G_{k\cdot}(v_{k\cdot})$.

[30]The main intuition is that after the redistribution of firms profits, agents that are receiving the transfer will have the following ex-post utility: $\widetilde{U}_{ij} = \ln u_{kj} + \beta_{kj} \ln\left\{ w_{kj}(1 + \phi(s, w; \lambda)) \right\} + \epsilon_{ij}$.

[31]This welfare function extends and generalizes the one considered in Lamadon et al. (2022) that assumes $\beta_{kj} = \beta, \mathcal{R} = (\mathcal{K} \times \mathcal{J})$, and full employment, i.e. $s_{k0} = 0$ for all $k \in \mathcal{K}$.

Convex duality implies the following relationship between the adjusted social surplus function and its convex conjugate:[32]

$$G_{k\cdot}(\widetilde{v}_{k\cdot}) = \sum_{j=0}^{J} \widetilde{v}_{kj} s_{kj} - G^*_{k\cdot}(s_{k\cdot}). \tag{4.3}$$

Using the above relationship (4.3), the welfare function becomes:

$$\mathcal{W}(\Xi, \lambda, \mathcal{R}) = \left[\sum_{(k,j)\in\mathcal{K}\times\mathcal{J}_0} m_k v_{kj} s_{kj} + \ln[1+\phi(s,w;\lambda,\mathcal{R})] \sum_{(k,j)\in R} m_k \beta_{kj} s_{kj} \right] - \sum_{k=1}^{K} m_k G^*_{k\cdot}(s_{k\cdot}). \tag{4.4}$$

where $\phi(s,w;\lambda,\mathcal{R}) = \frac{\sum_{j=1}^{J}\pi_j}{\sum_{(k,j)\in\mathcal{R}} w_{kj}\ell_{kj}}$.

The welfare function in eq. 4.4 is the summation of two main components: (i) a summation of the deterministic gains obtained in the equilibrium matching by all agents directly through their wages, preferences for amenities, and transfer of firms profits, and (ii) a measure of the randomness existing in the market. This last term is essentially due to the unobserved heterogeneity on the workers utilities. When ϵ follows the Logit distribution $\sigma_{kg} = 1$ for all (k, g), $-G^*_{k\cdot}(s_{k\cdot})$ is the usual Shannon entropy, which in information theory is considered as a natural measure of statistical disorder.[33] Following Galichon and Salanié (2022) we denote the generalized entropy as $-G^* \equiv \sum_{k=1}^{K} m_k G^*_{k\cdot}(s_{k\cdot})$. This captures the level of *incomplete information* in the market and allows us to construct a useful index of market concentration which is directly linked to the social welfare function. In the Nested Logit Economy, we consider the generalized exponential concentration index (GCI).[34]

$$\begin{aligned} GCI(s_{k\cdot}) \quad &\equiv \quad e^{G^*_{k\cdot}(s_{k\cdot})} = \exp\left[s_{k0}\ln s_{k0} + \sum_{g=1}^{G}\left[\frac{1}{\sigma_{kg}} \sum_{j\in N_g} s_{kj}\ln s_{kj} + (1-\frac{1}{\sigma_{kg}}) s_{kg}\ln s_{kg} \right] \right] \\ &= \quad \left[\Pi_{g=0}^{G} \left(\underbrace{\exp\left\{ \sum_{j\in N_g} s_{kj|g}\ln s_{kj|g} \right\}}_{\text{within-group concentration index}} \right)^{\frac{s_{kg}}{\sigma_{kg}}} \right] \times \underbrace{\left[\exp\left\{ \sum_{g=0}^{G} s_{kg}\ln s_{kg} \right\} \right]}_{\text{between-group concentration index}}, \end{aligned} \tag{4.5}$$

where $N_0 \equiv \{0\}$ and $\sigma_{k0} = 1$. To provide more intuition for this expression, consider the special case where we have symmetric firms in each local market, i.e. $v_{kjg} = v_{klg}$, for $j \neq l$, full employment, i.e. $w_{k0} = 0$, and the same correlation across nest, i.e., $\sigma_{kg} = \sigma_k$. In this case, the market shares in the Nested Logit Economy simplify to $s_{kjg} = \frac{1}{N_g} \times \frac{1}{G}$, and we

[32]Please see Galichon and Salanié (2022) for more detailed discussion.

[33]In their one-to-one matching model with perfect competition, Caldwell and Danieli (2024) make use of the continuous version of the Shannon entropy index as a measure of industrial concentration.

[34]Please see Allen and Rehbeck (2019), Example 7, for more details.

can write the GCI as:

$$GCI(s_{k\cdot}) = e^{\left\{-\frac{1}{\sigma_k G}\sum_{g=1}^{G}\ln N_g\right\}} \times e^{-\ln G}.$$

This simplified version of the GCI highlights that a stronger correlation of workers unobserved tastes for firms increases the within-group concentration and thus the overall concentration. In general, the GCI has a very natural and intuitive interpretation. It is a weighted function of "within group" concentration values, and a "between group" component.[35] As pointed out in Maasoumi and Slottje (2003), this type of decomposability of a concentration index is very useful for examining heterogeneity across different local markets. It allows one to identify areas with high concentration levels and those firms that contribute to concentration. It also allows policy makers to identify the impact of various changes and policy decisions on any desired group of firms and local markets, as well as on the overall concentration. It is worth noting that with this decomposability feature, any changes that increases "within group" concentration but keeping the "between group" concentration component constant will lead to an increase of the overall concentration. The widely-used Herfindahl index (HHI) does not have this decomposability feature.

Finally, we can explicitly link the social welfare to a market concentration index as follows:

$$\mathcal{W}(\Xi, \lambda, \mathcal{R}) = \left[\sum_{(k,j)\in\mathcal{K}\times\mathcal{J}_0} m_k v_{kj} s_{kj} + \ln[1+\phi(s, w; \lambda, \mathcal{R})] \sum_{(k,j)\in R} m_k \beta_{kj} s_{kj}\right] - \sum_{k=1}^{K} m_k \ln GCI(s_{k\cdot}). \quad (4.7)$$

This latter equation allows one to assess how changes in local concentration affect social welfare. It demonstrates that social welfare is a decreasing function of the GCI, holding fixed the deterministic gains from matching.

5. Econometric model: Identification and Estimation

In this section, we study identification of the structural parameters of the model when considering a Nested Logit Economy. Thus far, we have considered a static model. Here we assume that the econometrician has panel data linking workers to firms. We denote t the unit of time and let $t \in \{1, ..., T\}$. For the sake of tractability, we assume that both the econometrician and firms observe worker type k.[37] Provided that worker type k is known by the econometrician, our identification can be summarized in two steps. First, we

[35] The different components that form the CGI correspond to the Hannah-Kay (1971) concentration index for $\alpha = 1$. Indeed,

$$HK_\alpha(s_{k\cdot}) = \begin{cases} \left(\sum_j s_{kj}^{\alpha}\right)^{\frac{1}{\alpha-1}} \text{ if } \alpha > 0, \alpha \neq 1,^{36} \\ \exp\left(\sum_j s_{kj}\ln s_{kj}\right) \text{ if } \alpha = 1. \end{cases} \quad (4.6)$$

[37] If there are worker characteristics that influence firms' labor demand that are unobserved by the econometrician, we suggest employing the approach outlined by Bonhomme et al. (2019) to estimate these unobserved

identify the labor supply parameters using an instrumental variable approach. Second, we identify the production function parameters by exploiting firm optimization together with an instrumental variables strategy. It is worth noting that our identification approach does not require solving the model equilibrium, so identification is robust to the existence of multiple equilibria.

5.1. **Identifying the Labor Supply Parameters.** Consider the Nested Logit Economy where firms are partitioned into nests or, equivalently, local labor markets g. We define a firm's inside share, $s_{kj|gt}$, as the firm's employment share of worker type k in year t in labor market g. Following Berry (1994), we can derive the following quasi-supply function:

$$\ln \frac{s_{kjt}}{s_{k0t}} = \beta_k \ln \frac{w_{kjt}}{w_{k0t}} + (1 - 1/\sigma_{kg}) \ln s_{kj|gt} + \ln u_{kjt} \tag{5.1}$$

where s_{k0t} and w_{k0t} are the labor market share and earnings of non-employed workers of type k in period t, and u_{kjt} are the unobserved non-pecuniary benefits offered by firm j to workers of type k in year t. We restrict the labor supply parameters to be fixed over time and across firms, but allow σ_{kg} to vary by local market and labor type, and β_k to vary by labor type.

The parameters of interest are the distribution of unobserved amenities (u_{kjt}) and labor supply elasticities ($\mathcal{E}_{jkt}$) across all firms and worker types. The identification challenge in estimating equation 5.1 is that both the wage and the inside share are potentially correlated with the unobserved amenities and thus endogenous. The most common approach in the industrial organization literature, which we adopt here, is to identify the model parameters using instrumental variables (IV) for wages and the inside share. In particular, we follow the IV strategy developed by Lamadon et al. (2022): We rewrite our labor quasi-supply function (5.1) in changes as

$$\Delta_{e,e'} \ln \frac{s_{kjt}}{s_{k0t}} = \beta_k \Delta_{e,e'} \ln \frac{w_{kjt}}{w_{k0t}} + \tilde{\sigma}_{kg} \Delta_{e,e'} \ln s_{kj|gt} + \Delta_{e,e'} \ln u_{kjt} \tag{5.2}$$

where $\Delta_{e,e'} x_t \equiv x_{t+e} - x_{t-e'}$ and $\tilde{\sigma}_{kg} \equiv (1 - 1/\sigma_{kg})$. This provides a linear function of model parameters which can be consistently estimated using the two-stage least squares (2SLS) method under a relevance and exogeneity assumption.

For the instruments to be valid, we need them to be correlated with long changes ($e+e'+1$ periods) in the log wage ratio and log inside share (relevance), but orthogonal to long changes in amenities (exogeneity). To accomplish this, we use internal instruments relying on timing assumptions similar to Lamadon et al. (2022). Our instruments of choice are short (one-period) changes in the log establishment revenue ($\Delta \log R_{jt}$), the log inside share

characteristics. This requires an additional set of assumptions that must be carefully justified before implementation. The specifics of this methodology applied to our setting are provided in Appendix A.2.

($\Delta \log s_{kj|gt}$), and the log of the sum of the inside shares for all other labor types employed by the firm ($\Delta \log s_{\sim kj|gt}$). Short changes in these variables will be correlated with long changes in log wages and market shares as long as the labor productivity processes (defined as $\tilde{\gamma}_{kjt}$ in the next section) which determine them are sufficiently persistent.[38] These instruments satisfy the exogeneity assumption as long as the amenity process is sufficiently transitory. Lamadon et al. (2022) argue that unobserved firm-specific job amenity shocks are well approximated by a MA(1) process, showing that given this specification a choice of $e \geq 2$ and $e' \geq 3$ satisfies the exogeneity assumption. Here, we set $e = 2$, $e' = 3$ and assume that $\mathrm{Cov}(\tilde{\gamma}_{kjt+e} - \tilde{\gamma}_{kjt-e'}, \Delta z_{jkt}) \neq 0$ and $\mathrm{Cov}(\ln u_{kjt+e} - \ln u_{kjt-e'}, \Delta z_{jkt}) = 0$ for each $z_{jkt} \in \{\log R_{jt}, \log s_{kj|gt}, \log s_{\sim kj|gt}\}$, where $\Delta z_{jkt} \equiv z_{jkt} - z_{jkt-1}$. Importantly, this does not restrict correlations between the average *level* of firm-level amenities and labor productivity, nor does it preclude the firm from having chosen the overall level of amenities endogenously. In Appendix A.3.2, we provide further details on our estimation approach and give formal assumptions under which our estimation procedure provides consistent estimates for the labor supply parameters.[39] Given the estimated parameters, we can then use equation 5.1 to recover unobserved amenities ($\ln u_{kjt}$).

Finally, it is worth noting that our approach to identifying labor market power does not rely directly on the pass-through of firm-specific productivity shocks. The link between pass-through and labor market power is much more complicated in the presence of strategic interactions as shown above and therefore does not be used to identify the structural parameters.

5.2. **Identifying the Labor Demand Parameters.** We assume that the production function for firm j at time t takes the following form:

$$F_t^j(\ell_{\cdot j}) = \left(\sum_{k \in \mathcal{C}_t^j} \tilde{\gamma}_{kjt} \ell_{kjt}^{\rho_k} \right)^{\alpha_{jt}}, \tag{5.3}$$

where $\tilde{\gamma}_{kjt} = \theta_{jt} \gamma_{kjt}$ with $\sum_{k \in \mathcal{C}_t^j} \gamma_{kjt} = 1$. Recall that C_t^j is the set of worker types k employed by firm j in period t. With this specification, the first-order condition (FOC), i.e.

[38] In our results, we estimate the labor productivity process as an AR(1) and find that it is highly persistent.

[39] The industrial organization, trade, and labor literatures provide a number of possible instrumental variable strategies. In Appendix A.3.2 we discuss various instrumental variables that have been proposed in the IO literature such as "BLP" instruments (using the characteristics of competing firms in the market) and Hausman instruments (the tendency of firms to set correlated wages across establishments). We considered these instruments but found that they were not sufficiently strong in our setting. We also implemented a shift-share IV approach following Hummels et al. (2014) and Garin and Silvério (2023). We find labor supply parameter estimates that are comparable to our main estimates despite the fact that we are only able to construct the instrument for the small share of the firms in our sample who export.

Eq (2.7) becomes

$$\lambda_{jt}\alpha_{jt}\left(\sum_{k\in\mathcal{C}_t^j}\tilde{\gamma}_{kjt}\ell_{kjt}^{\rho_k}\right)^{\alpha_{jt}-1}\tilde{\gamma}_{kjt}\rho_k\ell_{kjt}^{\rho_k-1}=\frac{\mathcal{E}_{kjt}+1}{\mathcal{E}_{kjt}}w_{kjt} \tag{5.4}$$

Define $\tilde{w}_{kjt}\equiv\frac{\mathcal{E}_{kjt}+1}{\mathcal{E}_{kjt}}w_{kjt}$ (i.e.: the marginal revenue product of k-type labor at firm j) and take the ratio of the FOCs for different labor types $k,h\in\mathcal{C}_t^j$ to obtain:

$$\frac{\tilde{\gamma}_{kjt}\rho_k\ell_{kjt}^{\rho_k-1}}{\tilde{\gamma}_{hjt}\rho_h\ell_{hjt}^{\rho_h-1}}=\frac{\tilde{w}_{kjt}}{\tilde{w}_{hjt}} \tag{5.5}$$

Taking logs, we have the following log-linear equation:

$$\log\frac{\tilde{w}_{kjt}}{\tilde{w}_{hjt}}=(\rho_k-1)\log\ell_{kjt}-(\rho_h-1)\log\ell_{hjt}+\log\frac{\rho_k}{\rho_h}+\log\frac{\tilde{\gamma}_{kjt}}{\tilde{\gamma}_{hjt}}$$

with the last two terms being unobserved by the econometrician. The key parameters of interest are ρ_k, ρ_h, $\tilde{\gamma}_{kjt}$ and $\tilde{\gamma}_{hjt}$. The identification challenge is that both ℓ_{kjt} and ℓ_{hjt} may be correlated with the ratio $\frac{\tilde{\gamma}_{kjt}}{\tilde{\gamma}_{hjt}}$. However, with some assumptions on the structure of $\tilde{\gamma}_{kjt}$ we can obtain internal instruments which allows for consistent estimation of ρ_k and ρ_h. In particular, suppose that labor productivity for type k, $\tilde{\gamma}_{kjt}$, can be decomposed into an aggregate component $\bar{z}_{kt}$ and a firm-level component z_{jkt} such that $\tilde{\gamma}_{kjt}=\bar{z}_{kt}z_{jkt}$. Assume that the firm-level component follows an AR(1) process in logs: $\log z_{kjt}=\delta_k\log z_{kjt-1}+\bar{\varsigma}_k+\varsigma_{kjt}$ where ς_{kjt} is an i.i.d mean-zero innovation. Next, assume that the firm's choice of wages and labor are conditional on $\tilde{\gamma}_{.jt}$ and thus $\varsigma_{.jt}$, but that the innovation is independent from all lagged variables. Substitution leads to the following estimating equation, where we have assumed that $\delta_k=\delta_h\,\forall k,h$.

$$\begin{aligned}\log\frac{\tilde{w}_{kjt}}{\tilde{w}_{hjt}}=c_{kht}&+(\rho_k-1)\log\ell_{kjt}-(\rho_h-1)\log\ell_{hjt}+\delta\log\frac{\tilde{w}_{kjt-1}}{\tilde{w}_{hjt-1}}\\&-\delta(\rho_k-1)\log\ell_{kjt-1}+\delta(\rho_h-1)\log\ell_{hjt-1}+\varsigma_{khjt}\end{aligned} \tag{5.6}$$

where $c_{kht}\equiv\bar{\varsigma}_k-\bar{\varsigma}_h+(1-\delta)\log\frac{\rho_k}{\rho_h}+(\log\bar{z}_{kt}-\log\bar{z}_{ht})-\delta(\log\bar{z}_{kt-1}-\log\bar{z}_{ht-1})$ is a time-varying constant and $\varsigma_{khjt}\equiv\varsigma_{kjt}-\varsigma_{hjt}$ is i.i.d and mean zero. Note that ℓ_{kjt} and ℓ_{hjt} may be correlated with the error term ς_{khjt}. However, by assumption, ς_{khjt} is uncorrelated with lagged inputs, wages and revenues, allowing us to use functions of these lagged variables as instruments for contemporary input values.[40] This leads to identification of ρ_k, ρ_h and δ.

Estimating equation (5.6) is not straightforward to estimate as it is unclear how to choose the (k,h) pairs and construct the instruments/moments for each equation. To deal with this

[40] In the empirical application, we use functions (squares) of lags of the input price ratios and labor input quantities.

issue, we propose a multi-equation GMM approach which we discuss in detail in Appendix A.4.

Given a consistent estimator $\hat{\rho}_k$, we can rearrange the FOC (equation 5.5) to get

$$\tilde{\gamma}_{hjt} = A_{khjt}\tilde{\gamma}_{kjt} \tag{5.7}$$

where

$$A_{khjt} \equiv \frac{\tilde{w}_{kjt}^{-1}\ell_{kjt}^{\rho_k-1}\rho_k}{\tilde{w}_{hjt}^{-1}\ell_{hjt}^{\rho_h-1}\rho_h}$$

is a combination of data and known parameters. Recall that since $\tilde{\gamma}_{kjt} \equiv \theta_{jt}\gamma_{kjt}$ where $\sum_{k\in\mathcal{C}_t^j}\gamma_{kj} = 1$, we have

$$\sum_{h\in\mathcal{C}_t^j\setminus\{k\}}\gamma_{hjt} = \gamma_{kjt}\sum_{h\in\mathcal{C}_t^j\setminus\{k\}}A_{khjt} \Rightarrow (1-\gamma_{kjt}) = \gamma_{kjt}\sum_{h\in\mathcal{C}_t^j\setminus\{k\}}A_{khjt} \Rightarrow \gamma_{kjt} = \frac{1}{\sum_{h\in\mathcal{C}_t^j}A_{khjt}}$$

for all $k \in \mathcal{C}_t^j$. The first implication holds because $\sum_{k\in\mathcal{C}_t^j}\gamma_{kjt} = 1$, and the last one holds since $A_{kkjt} = 1$. This identifies γ_{kjt} for all k, j. So far, note that identification of γ_{kjt}, and ρ_k do not require any assumptions on the output market. To recover α_{jt} and θ_{jt} we assume perfect competition in output markets, meaning that each firm j is a price taker on the output market, i.e. $\lambda_{jt} = P_{jt}$ where P_{jt} is the exogenous price. Recall that from equation (5.4) we have:

$$\frac{\tilde{w}_{kjt}}{\gamma_{kjt}\rho_k\ell_{kjt}^{\rho_k-1}} = \lambda_{jt}\theta_{jt}^{\alpha_j}\alpha_{jt}\left(\sum_{k\in\mathcal{C}_t^j}\gamma_{kjt}\ell_{kjt}^{\rho_k}\right)^{\alpha_{jt}-1}.$$

Then, by re-arranging and noticing that at the optimum we have $\left(\sum_{k\in\mathcal{C}_t^j}\tilde{\gamma}_{kjt}\ell_{kjt}^{\rho_k}\right)^{\alpha_{jt}} = Q_j$ we obtain the following identification result:

$$\begin{aligned}\alpha_{jt} &= (\underbrace{\lambda_{jt}Q_{jt}}_{R_{jt}})^{-1} \times \underbrace{\frac{\mathcal{E}_{kjt}+1}{\mathcal{E}_{kjt}}w_{kjt}}_{\tilde{w}_{kjt}} \times \ell_{kjt} \times \frac{\sum_{k\in\mathcal{C}_t^j}\gamma_{kjt}\ell_{kjt}^{\rho_k}}{\gamma_{kjt}\rho_k\ell_{kjt}^{\rho_k}} \\ &= R_{jt}^{-1} \times \sum_{h\in\mathcal{C}_t^j}\tilde{w}_{hjt}\ell_{hjt}^{1-\rho_h}\rho_h^{-1} \times \sum_{k\in\mathcal{C}_t^j}\gamma_{kjt}\ell_{kjt}^{\rho_k},\end{aligned} \tag{5.8}$$

where $R_{it} \equiv P_{jt}Q_{jt}$ denotes firm j's total revenue. The second equality holds because of the following:

$$\frac{1}{\gamma_{kjt}} = \tilde{w}_{kjt}^{-1}\ell_{kjt}^{\rho_k-1}\rho_k \sum_{h\in\mathcal{C}_t^j}\frac{1}{\tilde{w}_{hjt}^{-1}\ell_{hjt}^{\rho_h-1}\rho_h} \quad\Rightarrow\quad \frac{\tilde{w}_{kjt}\times\ell_{kjt}}{\gamma_{kjt}\rho_k\ell_{kjt}^{\rho_k}} = \sum_{h\in\mathcal{C}_t^j}\tilde{w}_{hjt}\ell_{hjt}^{1-\rho_h}\rho_h^{-1}.$$

Finally, consider the identification of θ_{jt}. Recall that $R_{it} \equiv P_{jt}Q_{jt} = P_{jt}\theta_{jt}^{\alpha_{jt}}\left(\sum_{k\in\mathcal{C}_t^j}\gamma_{kjt}\ell_{kjt}^{\rho_k}\right)^{\alpha_{jt}}$. Therefore, we finally obtain $\tilde{\theta}_{jt}$ as

$$\tilde{\theta}_{jt} \equiv P_{jt}^{1/\alpha_{jt}}\theta_{jt} = \frac{R_{jt}^{1/\alpha_{jt}}}{\left(\sum_{k\in\mathcal{C}_t^j}\gamma_{kjt}\ell_{kjt}^{\rho_k}\right)}. \tag{5.9}$$

Note that we could recover θ_{jt} if we observe P_{jt} or normalize P_{jt} to 1.

6. Empirical Application

In this section, we apply our identification strategy to estimate the model parameters using population and firm administrative registers and linked employer-employee data from Denmark. Our estimation approach closely follows our identification arguments.

6.1. **Sample Construction and Descriptive Statistics.**

We use annual individual and firm registers and the linked employer-employee register IDA (Integrated Database for Labor Market Research) for the years 2001-2019. From the individual register, we get demographic and socio-economic worker characteristics and we identify unemployment and non-employment spells and income. From the firm register, we get yearly revenues arising from the firm's primary operation net of taxes and duties for private-sector firms. The linked employer-employee data contains information on salary, hours/days worked, industry, and workplace location of each employment contract every year. We combine the registers into a yearly panel dataset of workers through unique identifiers for individuals, firms, and establishments. We follow Taber and Vejlin (2020) and Berger et al. (2023) by focusing our empirical analysis on establishments which are linked to a physical location. Establishments are indexed by j and years by t. To get establishment-level revenue R_{jt}, we allocate firm revenue across establishments in proportion to their wage bills. Details on raw data, linkage of datasets, and construction of key variables are available in Appendix C.

We restrict the sample to all individuals between 26 and 60 years of age who work full-time as employees in the private sector and whose job is linked to a physical establishment. We exclude individuals employed in the public sector and the financial sector due to missing revenue data; financial sector firms are not legally required to report revenue and very few do. In total, our dataset consists of $12,742,746$ individual-year combinations. We assign individuals to 12 observable types k where each type is a combination of sex, age and education.[41]

For our empirical analysis, we collapse the dataset at the (k, j, t) level leading to $4,487,628$ observations. We further restrict this dataset to only establishments that have no missing values for any of our key variables. These include long and short changes in wages, market shares and inside shares for all other labor types employed by the establishments.[42] Our final dataset contains data for the years 2004-2017 and consists of $1,101,541$ observations at the (k, j, t) level.

We measure labor inputs in terms of full-time equivalents (FTE).[43] We use FTEs and worker-establishment linkages to calculate employment variables ℓ_{kjt}, s_{kjt}, and $s_{kj|gt}$ for each worker type k, in establishment j, in year t, overall and by market g. For every k, we also calculate the sum of the inside shares for all other labor types within the establishment, $s_{\sim kj|gt}$. We follow Taber and Vejlin (2020) by using non-employment (unemployment + non-participation) as the outside option. We calculate the share of non-employed workers in the economy every year by worker type k, s_{k0t}, by summing the non-employment spells at the k level and dividing by the total number of FTEs and non-employment spells in the data. The wage w_{ijt} for worker i at establishment j in year t is the total earnings for that worker in the year. We aggregate w_{ijt} to the (k, j, t) level by calculating the mean earnings w_{kjt} for each establishment j and each worker type k, in each year t. We also compute the mean non-employment income w_{k0t} for each worker type in the economy.[44]

We define a local labor market g as a commuting zone and industry pairing. We use the 5-digit industry classification based on the EU classification NACE Rev. 2 (Carré, 2008). After dropping the public and financial sectors, we have 15 industries. We use the commuting zones computed by Eckert et al. (2022) who use the Tolbert and Sizer (1996) method for

[41] Women represent 31.8 percent of the sample primarily due to women being overrepresented in the Danish public sector (which includes the education and health sectors).The full population of salaried jobs in Denmark in 2001-2019 is 49.3 percent female. This goes down to 35.8 percent when we drop the public sector and to further 31.8 percent when we exclude the financial sector and non-full-time jobs.

[42] In particular, for each variable x_{jkt}, we calculate short changes as $x_{jkt} - x_{jkt-1}$, and long changes as $x_{jkt+2} - x_{jkt-3}$, thus restricting the number of years available for the estimation to 2004-2017. Details are available in Appendix Table C.3.

[43] We calculate the full-time equivalent as the number of hours worked in the calendar year divided by the average number of full-time hours worked by full-time workers in Denmark over the same period, where we define a full-time worker as an individual who works 30+ hours a week. This implies that if an individual works as full-time in one establishment for six months, she will be counted as half of a FTE.

[44] Non-participation is defined as an individual not observed in the linked employer-employee data for a (part of the) year. Non-participation income is set to zero. Unemployment spells and unemployment income are observed directly in the data. Therefore, non-employment income consists of unemployment income for the unemployed workers. This includes cash assistance, unemployment benefits, leave benefits, and other assistance benefits, but—similarly to our measure of wages—it does not include long-term sickness or pension benefits.

Table 1. Worker distribution across k-groups, all years (2001-2019).

	k-group	share of worker-obs.	avg. earnings (in 2022 USD)	share of establishments
1	Female, 26-35, no college	0.046	50,775	0.177
2	Female, 26-35, college	0.033	64,750	0.092
3	Male, 26-35, no college	0.118	61,680	0.365
4	Male, 26-35, college	0.052	77,230	0.137
5	Female, 36-50, no college	0.110	57,347	0.298
6	Female, 36-50, college	0.052	79,674	0.122
7	Male, 36-50, no college	0.238	70,422	0.499
8	Male, 36-50, college	0.095	104,854	0.207
9	Female, 51-60, no college	0.059	56,847	0.192
10	Female, 51-60, college	0.018	77,465	0.054
11	Male, 51-60, no college	0.139	68,621	0.337
12	Male, 51-60, college	0.040	106,703	0.118
Number of worker-observations				12,742,746
Number of unique establishments				259,195

Average full-time equivalent (FTE) yearly earnings reported in real-2022 USD. The share of establishments refers to the share of establishments employing each k-group.

Denmark. Eckert et al. (2022) find 23 commuting zones in 2005. We drop six of the commuting zones that are small islands relatively separated from the mainland (Christiansœ, Bornholm, Samsœ, and Æro), and we merge the two North Jutland commuting zones of Aalborg and Frederikshavn. This leaves us with 16 commuting zones.

We display the 12 k-groups in Table 1, and report descriptive statistics for the full sample of workers based on the years 2001-2019. Column 1 reports the share of worker types in the sample, column 2 reports each k-group's average yearly earnings, and column 3 reports the share of establishments employing each k-group. The largest k-group is 36-50-year-old males with lower-than-college education, who make up 24 percent of the sample and are employed by half of the establishments. The smallest group is 51-60-year-old women with a college education, making up only 1.8 percent of the sample and employed by only 5.4 percent of the establishments. The highest earning group is 51-60-year-old males with a college education with average earnings of $106,703$ USD. The lowest earning group is 26-35-year-old females with lower-than-college education with average earnings of $50,775$ USD. The last column of Table 1 shows the share of establishments employing each k-group is between 5 and 50 percent, reflecting that the number of establishments which are truly available in the labor market for a particular type of worker is lower than the total number of establishments.

Appendix Tables D.1 and D.3 report establishment characteristics overall and by commuting zone and industry. Firms are composed of 1.2 establishments on average. This number is similar across commuting zones and industries. Each establishment employs on average 7.4 workers from 2.6 different k-groups, earns roughly 5.2 million USD in revenue, and pays an average wage of $59,000$ USD.[45]

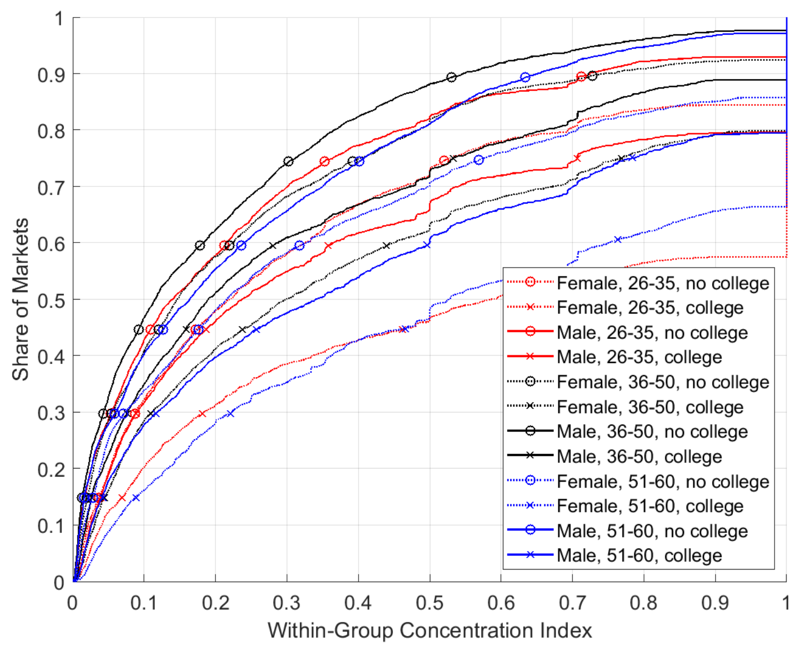
Commuting zones and industries (and therefore local markets) vary substantially in the number and type of establishments. The largest commuting zone is Copenhagen, containing around one third of all establishments in Denmark (over $80,000$ unique establishments over the sample period). Copenhagen also contains the largest establishments paying on average the highest wages. On the other hand, there are also very small commuting zones with under $2,000$ unique establishments during the time period 2001-2019 (i.e., Ribe and Thisted). In terms of industrial breakdown, the largest industry for number of establishments is wholesale and retail trade, followed by construction and knowledge-based services. Some industries such as mining, electricity, and water supply are quite small. Within each local labor market, there are 348 establishments on average (across years). However, the median number of establishments per market is 106 reflecting the skewness of this distribution.

6.2. **Empirical Analysis of the GCI.** Figure 1 shows the distribution of the within-group concentration index for each k-group (eq. 4.5) across local markets. As a reference point, if we assume that there are 5 symmetric establishments with an equal market share (which is usually interpreted as corresponding a moderate level of concentration). This corresponds to an HHI of 0.2 and a within-group concentration index of approximately 0.5. According to this benchmark, roughly 14 percent of local markets have a concentration level above 0.50 when averaging across k-groups (25 percent with the HHI). Moreover, the average level of concentration of these concentrated local markets is around 0.86.

The other notable feature of Figure 1 is the significant heterogeneity in concentration by worker type. Local markets for highly educated workers (both males and females) tend to be more concentrated than local markets for less educated workers and local markets for females are more concentrated than local markets for males (at all education levels). This can further be seen in Table 2 which aggregates the within-market concentration index across local markets according to equation (4.5) for each k-group. In particular, column 1 shows the overall GCI which is the product of the within-market index aggregated using a weighted geometric mean (column 2) and the between-market index (column 3). The rows of the table are sorted from the most concentrated to least concentrated according to the overall GCI. Non-college-educated females aged 51 to 60 are the group facing the highest market concentration, while college-educated males aged 36 to 50 are the group facing the least concentration.

[45]These statistics refer to the full sample, Appendix Tables D.2 and D.4 replicate the same statistics for the restricted estimation sample. The selection process leaves us with a subsample of establishments that are larger both in terms of size (11.6 workers and 3.5 k-groups) and revenue (8.9 million dollars).

FIGURE 1. Distribution of Within-Group Concentration Index by k-group



Cumulative distribution of local markets across the within-group component of the Generalized Concentration Index (WCI) from equation 4.5 (equivalent to the Entropy Index). WCI calculated on the full population of private sector establishments in Denmark (step 1 in appendix table C.3). Local market: Commuting Zone×Industry.

The decomposability property of the GCI allows one to have a better understanding of the source of the overall concentration index. This can be seen in Table 2 which shows that the low magnitude of the overall GCI (column 1) is driven more by the low level of between-local market concentration (column 3). On the other hand, column (2) shows that there is much more concentration within local markets than between local markets.

As a comparison to the overall GCI, in column (6) we report the overall HHI. The table shows that the ranking of concentration by k-group according to the HHI is very different than the ranking based on the GCI. This is not surprising since different concentration indices rely to varying degrees on different moments of the distribution of markets shares. The HHI captures only two moments of the distribution of the market shares (the mean and variance), while entropy-type concentration indices (like the GCI) additionally capture higher-order moments (Maasoumi and Theil, 1979). An additional factor that leads to differences between the GCI and the HHI is how it aggregates information across local markets. In particular, equation (4.5) shows that both the overall GCI (column 1) and the within-group GCI (column 2) depend on σ_{kg} which captures the degree of correlation of

Table 2. Generalized Concentration Index Across k-groups

	k-group	GCI Overall	GCI Within-Group	GCI Between-Group	Mean Local WCI	Mean Local HHI	Overall HHI
9	Female, 51-60, no college	0.056	0.623	0.089	0.1554	0.146	0.0001
1	Female, 26-35, no college	0.051	0.491	0.102	0.216	0.173	0.0000
10	Female, 51-60, college	0.047	0.591	0.078	0.350	0.286	0.0002
2	Female, 26-35, college	0.034	0.503	0.066	0.316	0.252	0.0003
5	Female, 36-50, no college	0.024	0.468	0.051	0.131	0.125	0.0001
6	Female, 36-50, college	0.018	0.410	0.042	0.238	0.211	0.0005
11	Male, 51-60, no college	0.013	0.454	0.028	0.079	0.099	0.0001
4	Male, 26-35, college	0.012	0.390	0.032	0.269	0.213	0.0005
3	Male, 26-35, no college	0.012	0.387	0.030	0.129	0.132	0.0001
12	Male, 51-60, college	0.011	0.395	0.027	0.227	0.208	0.0003
7	Male, 36-50, no college	0.009	0.384	0.023	0.068	0.086	0.0001
8	Male, 36-50, college	0.008	0.323	0.024	0.175	0.176	0.0005

Columns 1-3: Generalized Concentration Index (GCI) and the contribution of the within- and between-group components as in equation 4.5. Column 1 is the product of columns 2 and 3. Column 4: Within-group Concentration Index (WCI) as in equation 4.5, calculated as the arithmetic mean of the WCI computed for each local market. Column 5-6: local and overall Herfindahl-Hirschman Index. The local index is calculated as the arithmetic mean of the HHI computed for each local market. The overall HHI is calculated using the whole of Denmark as one market. We rank the k-groups from most concentrated to least concentrated according to the GCI. We calculate the GCI for the full population of private sector establishments in Denmark (step 1 in Appendix Table C.3), extrapolating the σ_{kg} estimates obtained with the restricted estimation sample (step 5 in Appendix Table C.3). All reported numbers are averages over the period 2001-2019.

worker preferences within the local market g. To see how this matters in practice, Table D.5 shows that low-educated females aged 51 to 60 have a relatively high σ_{kg} estimate. Even though this worker type has below-average concentration levels when ranked according to the HHI (both overall and within-market average, columns $5-6$), it is the most concentrated type when ranked according to the overall GCI, and this difference is driven by the within-group GCI which weights the local entropy index using σ_{kg} (column 2).

The decomposability property of the GCI allows us to identify the local markets that contribute the most to overall concentration by computing $\frac{1}{GCI(s_{k\cdot})}\left(\exp\left\{\sum_{j\in N_g} s_{kj|g}\ln s_{kj|g}\right\}\right)^{\frac{s_{kg}}{\sigma_{kg}}}$ for all local markets and ranking them from the highest to the lowest. In Denmark, the markets that contribute the most are a combination of mining and quarrying typically in smaller commuting zones (based on population counts). Electricity, gas and steam and Water supply/sewage are also large contributors. Construction in West-South Zeeland is a highly concentrated market for high-educated young women, Real Estate in North-West Jutland for young low-educated men, and accommodation and food services in North-West Jutland for low-educated older men.

6.3. **Estimates of Labor Supply.** We report moments of the distribution of the average labor supply elasticity and markdown estimates in Table 3. The average elasticity across

TABLE 3. Overview of Labor Supply Elasticities and Markdown Estimates.

Estimated Parameter		Mean	Median	P10	P90
Labor Supply Elasticity (eq. 3.1)	$\mathcal{E}_{kjt}$	5.790	5.429	2.800	8.665
Markdown $\left(\text{md}_{kj} = \frac{\mathcal{E}_{kj}}{1+\mathcal{E}_{kj}}\right)$	md_{kjt}	0.829	0.844	0.737	0.897
Cross-wage Super-elasticity (eq. 3.3)	ζ_{kjt}	-0.009	-0.002	-0.019	-0.000

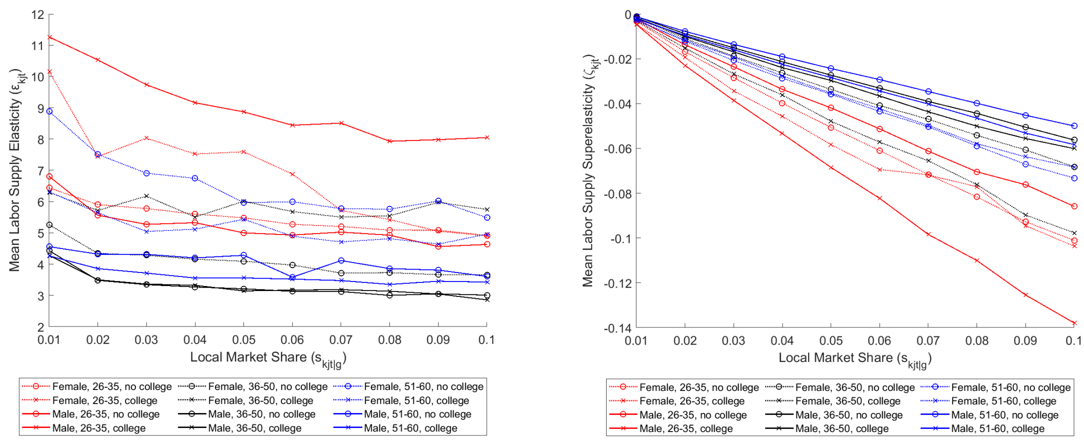
Estimated labor supply elasticities, markdowns, and cross-wage super-elasticities from the labor supply model. Moments of the estimated distributions of the establishment- and k-group-level elasticities and markdowns. We show the underlying estimates for β_k and σ_{kg} in Appendix Table D.5, and estimates for labor supply elasticities and markdowns separately by k-group in Appendix Table D.6.

all worker types, establishments, and years is 5.790, and the average markdown is 0.829, meaning that on average wages are marked down 17 percent relative to the marginal revenue product of labor.[46] There is significant heterogeneity in the distribution of labor supply elasticity across establishments and workers, with the 10th and 90th percentiles being 2.800 and 8.665, respectively. Appendix Table D.6 shows that the elasticities calculated using the IV-estimated parameters are larger than the corresponding OLS estimates. Appendix Table D.5 contains the underlying parameter estimates for each k-group. Columns (1) and (4) report the estimates for β_k and columns (2-3) and (5-6) report estimates for σ_{kg}. Our IV parameter estimates are reasonably well behaved with β_k estimates on average equal to 1.300 and σ_{kg} on average equal to 4.057. Our IV estimates for β_k are significantly larger than our OLS estimates implying significant downward bias in OLS. The IV estimates for σ_{kg} are slightly smaller than the corresponding OLS estimates.

Our estimate of the mean labor supply elasticity is comparable to existing estimates ranging between 3 and 5 (see Card, 2022, and references therein). In particular, Lamadon et al. (2022) estimate a labor supply elasticity of 4.2, and Kroft et al. (2023) find estimates ranging between 3.5 and 4.5 for the US construction sector. Berger et al. (2022a) estimate a distribution of firm-specific labor supply elasticities, the average across firms weighted by firm payroll is below 5 and the unweighted average across firms is above 9. The experimental literature finds a wider range of estimates between approximately 2 and 10 (Dube et al., 2018; Sokolova and Sorensen, 2021; Bassier et al., 2022; Emanuel and Harrington, 2022). A key feature of our framework is that elasticities vary by worker type, establishment, and market. To examine this heterogeneity, Figure 2 (a) displays labor supply elasticities by

[46]Note that a markdown of 0.829 is slightly lower than what one would obtain by computing the markdown using our average elasticity estimate of 5.790. This is because the markdown is a non-linear function of the elasticity implying that the average markdown does not equal the ratio of the average elasticity over 1 plus the average elasticity.

FIGURE 2. Labor supply elasticities by local market share and worker type



(A) labor supply elasticity ($\mathcal{E}_{kjt}$) (B) cross-wage super-elasticity (ζ_{kjt})

Local market: Commuting Zone×Industry. Panel (a) plots the estimated labor supply elasticities ($\mathcal{E}_{kjt}$) over the local market share ($s_{kjt|g}$), by k-group. Panel (b) plots the estimated labor supply super-elasticities (ζ_{kjt}) over the local market share ($s_{kjt|g}$), by k-group. Establishment-level elasticities are averaged across years and local markets by the establishment local market share bin (10 bins between 0 and 0.1).

worker type k as a function of the local market share $s_{kj|gt}$. We construct this figure by binning establishments according to their observed inside market share for each worker type and then taking an average of the labor supply elasticity across establishments and years. The average labor supply elasticity estimate masks significant heterogeneity, as mean elasticities across worker types range from 4.070 to 10.747. On the worker side, we see that younger workers tend to have significantly higher elasticities than older workers. For younger workers, more educated workers are more elastic than less educated workers but this pattern reverses for older workers. Younger women have similar or lower elasticity estimates than younger men, while women aged 36-50 and 51-60 have higher elasticities than men with the same age and education.[47]

On the establishment side, Figure 2 (a) shows that larger establishments—measured according to their inside share—face smaller elasticities and hence mark down wages further below the marginal revenue product of labor. Equation (3.1) shows that this relationship is not purely mechanically driven by the nested logit functional form since the elasticity also depends on the overall market share and on the k-group- and market-specific variance

[47]The experimental literature finds on average that women have lower labor supply elasticities than men (Sokolova and Sorensen, 2021), but there are a few exceptions. For example, using experimental evidence from Uber drivers in Houston, Caldwell and Oehlsen (2022) do not find any evidence that firm-specific elasticities differ by gender.

of the idiosyncratic amenities. Figure 2 (b) plots the cross-wage super-elasticities. For very small establishment, this is close to 0 and it declines as establishments get larger. Larger establishments have significantly more power to widen markdowns by decreasing employment relative to small establishments. These results also satisfy a key requirement for the uniqueness proof of the model equilibrium.

Given the labor supply estimates, we recover the establishment and k-group specific amenity terms u_{kjt} using equation (5.1). To investigate how deterministic preferences for amenities vary across job characteristics, we regress the log of the estimated u_{kjt} on fixed effects for commuting zone, industry, and several firm/establishment characteristics such as the log of size, wage bill, and revenue. We also control for k-group fixed effects, since the preferences terms u_{kjt} are identified up to a normalization for each k-group. We report these results in Table 4. Our estimates indicate that Copenhagen is the most desired commuting zone, with other large metro areas such as Aarhus and Odense ranked closely. The Knowledge-based services, Manufacturing and Transportation sectors have relatively high-value amenities, while Mining, Food services and Utilities have low-value average amenities. Finally, establishments with high-value amenities tend to have more workers, lower wages and lower revenue than low-value amenity establishments, conditional on industry and location. These results are in line with those for the US reported in Sorkin (2018), who finds evidence of low-value amenities for mining, construction, and transportation, as well as a strong contribution of establishment location to amenity values.

6.4. **Estimates of Labor Demand.** We report the production function estimates in Table 5. Panel A reports the average IV and OLS estimates of the labor substitution parameters ρ_k and the persistence of labor productivity δ. The IV estimates for ρ_k are 0.993 on average, range between $0.935 - 1.029$, are typically not statistically different from 1, and are fairly similar to the OLS estimates. These estimates imply that the different labor types in our context are highly substitutable, although the actual elasticity of substitution between two labor types at a given establishment will also depend on the relative employment/input levels of these two labor types, as we show below. We find that labor productivity is highly persistent with $\delta = 0.806$.

Panel B reports moments of the distribution of the establishment-level parameters. The distribution of α_{jt} is significantly skewed with a mean of 0.214 and median of 0.181. Similarly, the distribution of the overall productivity term $\tilde{\theta}_{jt}^{\alpha_{jt}}$ is highly skewed: the 90-10 ratio for private sector establishments in Denmark is 22.354.[48]

[48]This appears high relative to estimated firm productivity ratios in the industrial organization literature; however the measures should not be directly compared, as our model-relevant measure of productivity subsumes both TFP and firm variation in non-labor inputs (capital and intermediates/materials). These

Table 4. Correlation between Deterministic Preferences for Amenities and Establishment Characteristics

	$log(u_{kj})$	
Commuting zone (reference: North and East Zealand (Copenhagen))		
West and South Zealand (Slagelse)	-2.436	(0.002)
West and South Zealand (Køge)	-2.986	(0.003)
West and South Zealand (Nykøbing Falster)	-3.297	(0.004)
Fyn (Odense)	-1.818	(0.002)
Fyn (Svendborg)	-4.075	(0.004)
South Jutland (Sønderborg)	-2.517	(0.003)
South Jutland (Ribe)	-4.731	(0.006)
South Jutland (Kolding)	-1.793	(0.002)
Mid-South Jutland (Vejle)	-2.037	(0.002)
South-West Jutland (Esbjerg)	-2.744	(0.002)
West Jutland (Herning)	-2.431	(0.002)
North-West Jutland (Thisted)	-3.771	(0.005)
East Jutland (Aarhus)	-0.948	(0.001)
Mid-North Jutland (Viborg)	-3.030	(0.003)
North Jutland (Aalborg)	-1.330	(0.001)
Industry (reference: A. Agriculture, forestry, and fishery)		
B. Mining and quarrying	-1.406	(0.019)
C. Manufacturing	1.743	(0.004)
D. Electricity, gas, steam etc.	-0.624	(0.008)
E. Water supply, sewerage etc.	-0.754	(0.007)
F. Construction	0.880	(0.004)
G. Wholesale and retail trade	1.650	(0.004)
H. Transportation	0.942	(0.004)
I. Accommodation and food services	-0.189	(0.004)
J. Information and communication	0.967	(0.004)
L. Real estate	-0.146	(0.004)
M. Knowledge-based services	1.170	(0.004)
N. Travel agent, cleaning etc.	0.336	(0.004)
R. Arts, entertainment, recreation	-0.197	(0.006)
S. Other services	0.079	(0.006)
Log of establishment size (number of workers)	0.988	(0.002)
Log of establishment wagebill (thousands 2022 USD)	-0.909	(0.002)
Log of establishment revenue (thousands 2022 USD)	0.004	(0.001)
Log of firm size (number of workers)	-0.002	(0.000)
Observations	2,332,047	
R^2	0.866	

OLS of $log(u_{kj})$ on k-group, commuting zone, industry, and year indicators, and establishment characteristics (logarithm of firm and establishment size in number of workers, and logarithm of establishment wage bill and revenue). We report coefficients for commuting zone, industry, and establishment characteristics. Robust standard errors in parentheses.

We next use the production function and labor supply estimates to construct establishment j and k-group specific labor demand elasticities, η_{kjt}, which we report in Table 5, Panel C and, by k-group, in Appendix Table D.8. The labor demand elasticities are negative as expected (since increased wages decrease demand for each type of labor). The distribution

estimates are also usually reported within the manufacturing sector, whereas we consider the entire private sector.

Table 5. Overview of Labor Demand Parameter Estimates.

Panel A. Estimated Parameters from eq. (5.6)					
		IV		OLS	
Persistence of Labor Productivity	δ	0.806 [0.804; 0.808]		0.806 [0.803; 0.808]	
Labor Substitution Parameters (average of ρ_1–ρ_{12})	ρ_k	0.993 [0.980; 1.007]		0.992 [0.988; 0.997]	
Panel B. Distribution of Other Estimated Parameters					
		Mean	Median	P10	P90
Labor Productivity (eq. 5.7)	γ_{kjt}	0.288	0.224	0.087	0.541
Scale Parameters (eq. 5.8)	α_{jt}	0.214	0.181	0.059	0.417
TFP (eq. 5.9)	$\log(\tilde{\theta}_{jt}^{\alpha_{jt}})$	3.826	3.240	2.678	4.027
Panel C. Distribution of Labor Demand Elasticities and TFP Passthrough Lower Bound					
		Mean	Median	P10	P90
Labor Demand Elasticities (eq. 3.5)	η_{kjt}	-10.130	-5.317	-27.258	-1.797
TFP Passthrough Lower Bound (eq. 3.6)	$\min\left(\frac{\tilde{\theta}_l}{w_{kl}}\frac{\partial w_{kl}}{\partial \tilde{\theta}_l}\right)$	0.552	0.538	0.268	0.853

Parameter estimates for the production function in eq. (5.3). Panel A: we report the IV and OLS estimate of δ and the average of our estimates for ρ_k. Bootstrapped 95% confidence intervals in square brackets (Hall, 1992) (average of the 12 confidence intervals for ρ_k). The underlying parameter estimates are in Appendix Table D.7.
Panel B: moments of the estimated distributions of the establishment-level production function parameters (γ_{kjt}, α_{jt}, $\tilde{\theta}_{jt}^{\alpha_{jt}}$). Find the full distributions of α_{jt} and $\tilde{\theta}_{jt}^{\alpha_{jt}}$ in Appendix Figure D.1. We report the distribution of γ_{kjt} by k-group in Appendix Figure D.2.
Panel C: moments of the establishment-level labor demand elasticities (η_{kjt}) and the lower bound of the passthrough of TFP shocks to wages. We report the distribution of η_{kjt} by k-group in Appendix Figure D.8

is fairly skewed, with an average of -10.130 and a median of -5.317, which implies a 1 percent increase in wage decreases average labor demand by 5.317 percent. The distribution also masks significant heterogeneity, with median labor demand elasticities ranging from -2.961 for middle-aged males with no college degree, up to -12 for middle-aged and older females with a college degree.

Recall that the labor productivity parameters γ_{kjt} are normalized at the establishment level. Thus, estimates of γ_{kjt} only have a meaningful interpretation within establishments. To interpret relative differences in labor productivities across k-groups, we regress γ_{kjt} on establishment$\times$year and worker type fixed effects. The resulting worker type fixed effects are reported in Table 6. Generally, the estimates show that more educated workers have

TABLE 6. Within-establishment Heterogeneity in γ_{kjt} by k-group

	γ_{kjt}
k-groups (reference: Male, 36-50, college)	
Female, 26-35, no college	-0.141
Female, 26-35, college	-0.133
Male, 26-35, no college	-0.110
Male, 26-35, college	-0.112
Female, 36-50, no college	-0.119
Female, 36-50, college	-0.096
Male, 36-50, no college	-0.046
Female, 51-60, no college	-0.126
Female, 51-60, college	-0.084
Male, 51-60, no college	-0.057
Male, 51-60, college	-0.016
Constant	0.333
Observations	2,212,859
R^2	0.866

Estimates from OLS of γ_{kjt} on k-group fixed effect and year×establishment fixed effects (not reported). Robust standard errors all below 0.0005, $p < 0.001$.

higher productivity than less educated workers. We also see that younger workers (age 26-35) are less productive than older workers (age 36+).

To get a better sense of what our production function estimates imply for labor substitutability, we compute the Morishima elasticity of substitution (MEOS, Morishima (1967)). For the standard CES case with two inputs, the MEOS is equivalent to the standard Allen-Uzawa elasticity of substitution. However, when considering non-homogeneous production functions (such as ours), or production functions with 3 or more inputs (again such as ours), the MEOS more accurately represents the underlying substitution elasticities faced by the firm (Blackorby and Russell, 1989). An added difficulty in calculating the Allen-Uzawa elasticity of substitution in our setting is that firms have monopsony power in input markets. This means that it's unclear how to interpret formulations of the elasticity which rely on derivatives with respect to wages (since wages are chosen by firms and are not exogenous). To resolve this, we use the generalized MEOS derived by Kuga and Murota (1972), where the MEOS of input factor k by h is defined as:

$$MEOS_{khjt} = \frac{F^j_{ht}}{\ell_{kjt}} \frac{H_{khjt}}{H_{jt}} - \frac{F^j_{ht}}{\ell_{hjt}} \frac{H_{hhjt}}{H_{jt}}$$

where $F^j_{ht} = \partial F^j_t / \partial \ell_{hjt}$, ℓ_{kjt} is the level of labor input k, H_{jt} is the bordered Hessian for the production function for establishment j in period t, and H_{khjt} is the cofactor of the $\partial^2 F^j_t / \partial \ell_{kjt} \partial \ell_{hjt}$ term in H. We calculate the MEOS for every input pair, across every establishment, in every period, and report the mean input pair-specific elasticities in Table D.9. The estimated elasticities are quite high. Note that the MEOS is not symmetric, unlike the Allen-Uzawa elasticity of substitution. For example, the elasticity of substitution of low educated by high educated middle-aged males is 17, with the reverse being 43. The pattern of the average MEOS terms broadly follow the estimated ρ_k parameters, with young college educated females (k-group 2) having both the highest ρ_k parameter, and the lowest overall substitution elasticities. Similarly, middle-aged college educated males (k-group 8) have the lowest ρ_k and highest average substitution elasticities.

7. Counterfactual analyses

7.1. **Model-based Variance Decomposition.** In this section, we study the relative contributions of the different mechanisms of our model – heterogeneity in labor supply (deterministic and stochastic preferences for amenities) and heterogeneity in labor demand (heterogeneity in worker skill and production technology) on equilibrium wage inequality, concentration, and welfare. We follow the decomposition approach of Taber and Vejlin (2020) by sequentially eliminating each source of wage inequality and counterfactually predicting the effect on equilibrium wages. An important difference is that our counterfactual exercises take into account general equilibrium effects while Taber and Vejlin (2020) consider only partial equilibrium effects.

To fix ideas, recall that Ξ is the vector of model parameters, and let $\hat{\Xi}$ denote the empirical estimates of these parameters. We denote by $\mathbb{V}^p(\hat{\Xi})$ the variance of log wages predicted by our model. We obtain $\mathbb{V}^p(\hat{\Xi})$ by fixing $\hat{\Xi}$, solving the model equilibrium in eq. (3.4) using the Jacobi/Gauss-Seidel algorithm, and computing the variance of log wages. Our counterfactual analyses follows this approach by fixing the model parameters at some counterfactual values, i.e. Ξ^c, and then using the model equilibrium to compute the variance and other statistics associated with this counterfactual scenario, i.e. $\mathbb{V}^p(\Xi^c)$.

In each counterfactual scenario, we sequentially eliminate each source of wage inequality. Consider, for instance, the following parameters: $u_{kj}, \gamma_{kj}, \theta_j$. To study the contribution to wage inequality coming from Roy sorting and heterogeneity in the deterministic preference for amenities, we compute: $\mathbb{V}^p(\hat{\gamma}_{kj}, \hat{u}_{kj}, \hat{\theta}_j)$, $\mathbb{V}^p(\gamma_k^c, \hat{u}_{kj}, \hat{\theta}_j)$, and $\mathbb{V}^p(\gamma_k^c, u^c, \hat{\theta}_j)$. The ordering with which we shut down each mechanism is important because there are nonlinear interactions across mechanisms. We therefore implement a range of scenarios in which we shut down mechanisms in different orders:

[AB]: Heterogeneity in labor supply:

- [A]: Heterogeneity in the deterministic preferences for amenities: $u_{kj}^c = \overline{u}$.
- [B]: Heterogeneity in the stochastic preferences for amenities: $\beta_k^c = \overline{\beta}, \sigma_{gk}^c = \overline{\sigma}$.

[CD]: Heterogeneity in labor demand:

[C]: Worker skill heterogeneity within firm j: $\gamma_{kj}^c = \overline{\gamma}$, and no heterogeneity in the marginal rate of substitution between different types of workers: $\rho_k^c = \overline{\rho}$.

[D]: Heterogeneity in production technology: $(\tilde{\theta}_j^{\alpha_j})^c = \overline{\tilde{\theta}^{\alpha}}$, $\alpha_j^c = \overline{\alpha}$.

[E]: Constant return to scale in the production function: $\alpha_j^c = 1$ for all j.

The convention $\overline{X}$ is used to denote the observation-weighted mean of X, except for scenario D where we use the median due to the skewness of the production technology distribution. Although our primary focus is on the variance of log wages, we also report market concentration and welfare for each counterfactual scenario.

7.2. **Counterfactual Results.** Before presenting the results from each scenario, it is useful to highlight the key forces in the model that affect wages and how these interact. Equation 5.4 shows wages depend directly on α_{jt}, $\tilde{\gamma}_{kjt}$, ρ_k, the composite term $\left(\sum_{k\in\mathcal{C}_t^j} \tilde{\gamma}_{kjt}\ell_{kjt}^{\rho_k}\right)^{\alpha_{jt}-1}$, and the markdown. The first two channels jointly represent the marginal revenue product of labor (MRPL). First, consider the deterministic preference for amenities u_{kj}. The primary channel through which this mechanism affects wages is via the labor supplies ℓ_{kjt} in the composite term (and hence the MRPL) and the shares that enter the markdown. Second, consider the stochastic preference for amenities parameters β_k and σ_{gk}. These enter through the labor supply elasticities and hence the markdown and also indirectly through the endogenous labor supplies which enter the MRPL. Third, consider worker skill γ_{kj}, ρ_k and the heterogeneous production technology parameters, $(\tilde{\theta}_j^{\alpha_j})$ and α_j. These primitives affect the MRPL in two ways, one direct and another indirect through the composite term. These primitives also affect markdowns through the endogenous market shares. We will see that in general, the effects of shutting down heterogeneity in the model primitives on wages will depend on whether there is heterogeneity in the labor allocation (or worker skill) across firms through the composite term. Since this depends on whether there is heterogeneity in the deterministic preference for amenities (or worker skill), the forces in the model tend to interact.

At a broad level, each force can affect wages through both the markdown and the MRPL via the endogenous labor supplies and market shares. However, a key lesson that emerges from our counterfactual exercises is that heterogeneity in MRPL across firms swamps heterogeneity in the markdowns. This can be seen in Table D.10 which shows that the markdown accounts for very little of the variation in wages suggesting that this channel is relatively unimportant in accounting for overall wage inequality as compared to variation in the MRPL. Of course, this does not imply that markdowns are irrelevant for wage inequality; there must be a markdown on average for firm differences in the MRPL to matter for wage inequality across individuals.

Table 7 displays the results of our counterfactuals. The first counterfactual labeled "Truth" is calculated by solving the model using our estimated parameters. Reassuringly, the counterfactual log wage variance evaluated using the estimated parameters (0.1285) matches the empirical wage variance (0.1215) almost exactly. Thus, the estimated structural model is well suited to investigating the sources of wage inequality in Denmark.

Examining the next set of counterfactual exercises reveals the role of various mechanisms for understanding wage inequality in Denmark. First, across all scenarios, removing heterogeneity in the deterministic preferences for amenities ([A]) increases the variance of log wages by a factor of roughly two to four. Intuitively, this is because the estimated preference for amenities is positively correlated with production technology and all else equal, more productive firms pay higher wages. Thus, removing heterogeneity in the preference for amenities reduces labor supply at high productivity firms and increases wages while increasing labor supply at low productivity firms leading to lower wages.

Next, removing heterogeneity in the variance of the stochastic preference for amenities across worker types ([B]) has an ambiguous effect overall on wage inequality. When the deterministic preferences for amenities are restricted to be homogeneous (Scenarios 1 - 5), there is a reduction in wage dispersion of around $22\% - 29\%$. On the other hand, when the deterministic preferences for amenities are heterogeneous (Scenario 6 and 7), wage dispersion increases by roughly 20%. This ambiguity comes from the interactions emphasized above. In the former case, the reduction in preference dispersion mainly acts to reduce heterogeneity in the markdown through the labor supply elasticities and thus reduces wage inequality. In the latter case, there is an additional effect that operates through the allocation of labor across firms in the composite term which works in the opposite direction. We see that in this case, this effect via the MRPL dominates the direct effect which enters through the markdown.

We next examine the contribution of worker skill ([C]) and production technology ([D]) heterogeneity. On one hand, we find that eliminating heterogeneity in worker skill decreases the variance of log wages in all scenarios by roughly 30% to 50%. On the other hand, restricting heterogeneity in production technology has a more mixed effect. In Scenarios 3, 5 and 7, wage inequality increases by roughly 6% to 50% whereas in the other scenarios, it decreases by approximately 8% to 27%. The main difference between these scenarios is whether workers have heterogeneous deterministic preferences over amenities (Scenarios 3, 5 and 7) or they do not (Scenarios 1, 2, 4 and 6). In Scenarios 3, 5 and 7, the composite term is operative whereas in the other scenarios, it is not and only the direct effect dominates through the MRPL.

Finally, after removing the main sources of heterogeneity ([A], [B], [C], [D]), there is residual wage dispersion as can be seen in the second-last column of Table D.10. This is due to differences in demand for worker-types across firms along the extensive margin (i.e., heterogeneity in C_j) and differences in the underlying supply of different worker types (m_k). These both show up through employment, and drive variation in wages via the composite term. Removing curvature in the production function (setting $\alpha_j = 1$) drives this residual wage dispersion to zero.

Taken together, these results highlight that all the primary channels in our model drive wage inequality. Some mechanisms always increase inequality (heterogeneity in worker skill) while others always decrease inequality (heterogeneity in the deterministic preferences for amenities). In other cases (heterogeneity in the stochastic preferences for amenities and production technology), the direction of the effect on inequality depends on which other mechanisms are active in the model. These interaction effects are primarily due to the presence of decreasing returns to scale in the production function. In the presence of interaction effects, the order of the decomposition matters.

The other outcomes in Table 7 are concentration and welfare. A general lesson that emerges is that in most cases, concentration and social welfare are inversely related. Counterfactuals that lead to reductions in concentration tend to be associated with increases in welfare and vice-versa. One exception to this is when we eliminate heterogeneity in production technology ([D]). In this case, we see that both concentration and welfare increase. The main reason for this is that in this scenario, overall firm profits increase thus pushing welfare up.

8. Conclusion

This paper builds, identifies and estimates a structural two-sided matching model of the labor market featuring imperfect competition and rich heterogeneity. Our approach to studying market power in the labor market follows the modern dominant empirical Industrial Organization paradigm by developing a theory that is tied to the market, combined with a clear analysis of endogeneity, identification and instruments. We demonstrate identification of labor supply and demand parameters using instrumental variables and we estimate the model parameters using matched employee-employer data from Denmark covering the period 2001-2018. Our empirical results indicate heterogeneity in local markets according to concentration levels and market power of firms which vary both by worker characteristics and firm characteristics. We use our estimated structural model to shed light on the sources of wage inequality in Denmark. Our results indicate that some mechanisms always increase inequality (heterogeneity in worker skill) while others always decrease inequality

Table 7. Counterfactual Wage Dispersion, Concentration, and Welfare

Scenario 1						
Counterfactual Exercise:	Truth	A ($\overline{u}$)	B ($\overline{\beta}, \overline{\sigma}$)	C ($\overline{\gamma}, \overline{\rho}$)	D ($\overline{\theta}, \overline{\alpha}$)	E ($\alpha_j = 1$)
Variance of Log Wages	0.1285	0.427	0.3346	0.2764	0.2031	0.0005
Concentration (GCI)	0.0284	0.0096	0.0142	0.0102	0.0164	0.0028
Welfare	3.205	3.383	3.180	3.185	3.382	3.751
Scenario 2						
Counterfactual Exercise:	Truth	C ($\overline{\gamma}, \overline{\rho}$)	A ($\overline{u}$)	B ($\overline{\beta}, \overline{\sigma}$)	D ($\overline{\theta}, \overline{\alpha}$)	E ($\alpha_j = 1$)
Variance of Log Wages	0.1285	0.086	0.3876	0.2764	0.2031	0.0005
Concentration (GCI)	0.0284	0.0178	0.0084	0.0102	0.0164	0.0028
Welfare	3.205	3.233	3.415	3.185	3.382	3.751
Scenario 3						
Counterfactual Exercise:	Truth	C ($\overline{\gamma}, \overline{\rho}$)	D ($\overline{\theta}, \overline{\alpha}$)	A ($\overline{u}$)	B ($\overline{\beta}, \overline{\sigma}$)	E ($\alpha_j = 1$)
Variance of Log Wages	0.1285	0.086	0.0912	0.2827	0.2031	0.0005
Concentration (GCI)	0.0284	0.0178	0.0289	0.0118	0.0164	0.0028
Welfare	3.205	3.233	3.323	3.570	3.382	3.751
Scenario 4						
Counterfactual Exercise:	Truth	A ($\overline{u}$)	B ($\overline{\beta}, \overline{\sigma}$)	D ($\overline{\theta}, \overline{\alpha}$)	C ($\overline{\gamma}, \overline{\rho}$)	E ($\alpha_j = 1$)
Variance of Log Wages	0.1285	0.427	0.3346	0.3087	0.2031	0.0005
Concentration (GCI)	0.0284	0.0096	0.0142	0.0192	0.0164	0.0028
Welfare	3.205	3.383	3.180	3.428	3.382	3.751
Scenario 5						
Counterfactual Exercise:	Truth	D ($\overline{\theta}, \overline{\alpha}$)	C ($\overline{\gamma}, \overline{\rho}$)	A ($\overline{u}$)	B ($\overline{\beta}, \overline{\sigma}$)	E ($\alpha_j = 1$)
Variance of Log Wages	0.1285	0.1908	0.0912	0.2827	0.2031	0.0005
Concentration (GCI)	0.0284	0.0384	0.0289	0.0118	0.0164	0.0028
Total Welfare	3.205	3.307	3.323	3.570	3.382	3.751
Scenario 6						
Counterfactual Exercise:	Truth	B ($\overline{\beta}, \overline{\sigma}$)	A ($\overline{u}$)	D ($\overline{\theta}, \overline{\alpha}$)	C ($\overline{\gamma}, \overline{\rho}$)	E ($\alpha_j = 1$)
Variance of Log Wages	0.1285	0.1573	0.3346	0.3087	0.2031	0.0005
Concentration (GCI)	0.0284	0.091	0.0142	0.0192	0.0164	0.0028
Welfare	3.205	3.300	3.180	3.428	3.382	3.751
Scenario 7						
Counterfactual Exercise:	Truth	D ($\overline{\theta}, \overline{\alpha}$)	C ($\overline{\gamma}, \overline{\rho}$)	B ($\overline{\beta}, \overline{\sigma}$)	A ($\overline{u}$)	E ($\alpha_j = 1$)
Variance of Log Wages	0.1285	0.1908	0.0912	0.1101	0.2031	0.0005
Concentration (GCI)	0.0284	0.0384	0.0289	0.077	0.0164	0.0028
Welfare	3.205	3.307	3.323	3.397	3.382	3.751

Counterfactual estimates of log wage variance, concentration (GCI) and welfare for 7 different decomposition scenarios. In each scenario, each column represents a cumulative counterfactual exercise, where the effect is inclusive of previous columns. For example, Scenario 1 column 3 includes both exercise A and B and Column 4 includes exercises A, B and C. Concentration is the mean GCI across k-groups. Welfare is the social welfare function as in equation 4.7 in millions. Exercise A sets $u_{jk} = \overline{u}$, B sets $\beta_k = \overline{\beta}$ and $\sigma_{gk} = \overline{\sigma}$, C sets $\gamma_{kj} = \overline{\gamma}$ and $\rho_k = \overline{\rho}$, D sets $\theta_j^{\alpha_j} = \overline{\theta^{\alpha}}$ and $\alpha_j = \overline{\alpha}$, and E sets $\alpha_j = 1$. The overline represents the observation-weighted mean, except in D where it is the median.

(heterogeneity in the deterministic preferences for amenities). In other cases (heterogeneity in the stochastic preferences for amenities and production technology), the direction of the effect on inequality depends on which other mechanisms are active in the model. These interaction effects are primarily due to the presence of decreasing returns to scale in the production function.

Our framework can be used as a tool to study other sources of wage heterogeneity beyond overall inequality. For example, one could use it to examine the sources of wage gaps across groups (e.g., gender, race, or immigrant). Furthermore, one can use our framework to understand how mergers (as in Arnold, 2019; Prager and Schmitt, 2021) and labor market institutions such as unions (as in Dodini et al., 2022) and minimum wages (as in Berger et al., 2022b) affect market power, concentration and wage inequality.

References

Allen, Roy and John Rehbeck, "Identification with additively separable heterogeneity," *Econometrica*, 2019, *87* (3), 1021–1054.

Arnold, David, "Mergers and acquisitions, local labor market concentration, and worker outcomes," Technical Report 2019.

Azar, José A, Steven T Berry, and Ioana Marinescu, "Estimating labor market power," Technical Report, National Bureau of Economic Research 2022.

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

Azevedo, Eduardo M, "Imperfect competition in two-sided matching markets," *Games and Economic Behavior*, 2014, *83*, 207–223.

Bagger, Jesper and Rasmus Lentz, "An empirical model of wage dispersion with sorting," *The Review of Economic Studies*, 2019, *86* (1), 153–190.

Bassier, Ihsaan, Arindrajit Dube, and Suresh Naidu, "Monopsony in movers: The elasticity of labor supply to firm wage policies," *Journal of Human Resources*, 2022, *57* (S), S50–s86.

Benmelech, Efraim, Nittai K Bergman, and Hyunseob Kim, "Strong employers and weak employees: How does employer concentration affect wages?," *Journal of Human Resources*, 2022, *57* (S), S200–S250.

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

Berger, David W, Kyle F Herkenhoff, and Simon Mongey, "Minimum wages, efficiency and welfare," Technical Report, National Bureau of Economic Research 2022.

———, ———, Andreas R Kostøl, and Simon Mongey, "An Anatomy of Monopsony: Search Frictions, Amenities and Bargaining in Concentrated Markets," Technical Report, National Bureau of Economic Research 2023.

Berry, Steven T, "Estimating discrete-choice models of product differentiation," *The RAND Journal of Economics*, 1994, pp. 242–262.

Bhaskar, Venkataraman, Alan Manning, and Ted To, "Oligopsony and monopsonistic competition in labor markets," *Journal of Economic Perspectives*, 2002, *16* (2), 155–174.

Blackorby, Charles and R Robert Russell, "Will the real elasticity of substitution please stand up?(A comparison of the Allen/Uzawa and Morishima elasticities)," *American Economic Review*, 1989, *79* (4), 882–888.

Boal, William M and Michael R Ransom, "Monopsony in the labor market," *Journal of Economic Literature*, 1997, *35* (1), 86–112.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa, "A distributional framework for matched employer employee data," *Econometrica*, 2019, *87* (3), 699–739.

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

Caldwell, Sydnee and Emily Oehlsen, "Gender, outside options, and labor supply: Experimental evidence from the gig economy," Technical Report, Working Paper 2022.

——— and Oren Danieli, "Outside options in the labor market," *Review of Economic Studies*, 2024, p. rdae006.

Card, David, "Who set your wage?," *American Economic Review*, 2022, *112* (4), 1075–1090.

———, Ana Rute Cardoso, Joerg Heining, and Patrick Kline, "Firms and Labor Market Inequality:

Evidence and Some Theory," *Journal of Labor Economics*, January 2018, *36*, S13–S70.

Carré, H, "NACE Rev. 2, Statistical classification of economic activities in the European Community," *General and regional statistics*, 2008.

de Melo, Rafael Lopes, "Firm wage differentials and labor market sorting: Reconciling theory and evidence," *Journal of Political Economy*, 2018, *126* (1), 313–346.

Dodini, Samuel, Kjell G Salvanes, and Alexander Willén, *The dynamics of power in labor markets: Monopolistic unions versus monopsonistic employers*, IZA-Institute of Labor Economics, 2022.

Doraszelski, Ulrich and Jordi Jaumandreu, "Measuring the bias of technological change," *Journal of Political Economy*, 2018, *126* (3), 1027–1084.

Dube, Arindrajit, Alan Manning, and Suresh Naidu, "Monopsony and employer mis-optimization explain why wages bunch at round numbers," Technical Report, National Bureau of Economic Research 2018.

______, Jeff Jacobs, Suresh Naidu, and Siddharth Suri, "Monopsony in online labor markets," *American Economic Review: Insights*, 2020, *2* (1), 33–46.

Dupuy, Arnaud, Alfred Galichon, Sonia Jaffe, and Scott Duke Kominers, "Taxation in matching markets," *International Economic Review*, 2020, *61* (4), 1591–1634.

______ and ______, "A note on the estimation of job amenities and labor productivity," *Quantitative Economics*, 2022, *13* (1), 153–177.

Eckert, Fabian, Mads Hejlesen, and Conor Walsh, "The return to big-city experience: Evidence from refugees in Denmark," *Journal of Urban Economics*, 2022, *130*, 103454.

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu, "How costly are markups?," *Journal of Political Economy*, 2023, *131* (7), 000–000.

Eeckhout, Jan and Philipp Kircher, "Assortative Matching With Large Firms," *Econometrica*, 2018, *86* (1), 85–132.

Emanuel, Natalia and Emma Harrington, "Firm frictions and the payoffs of higher pay: Labor supply and productivity responses to a voluntary firm minimum wage," Technical Report, Working Paper 2022.

Galichon, Alfred and Bernard Salanié, "Cupid's invisible hand: Social surplus and identification in matching models," *The Review of Economic Studies*, 2022, *89* (5), 2600–2629.

Gandhi, Amit, Salvador Navarro, and David A Rivers, "On the identification of gross output production functions," *Journal of Political Economy*, 2020, *128* (8), 2973–3016.

Garin, Andrew and Filipe Silvério, "How Responsive Are Wages to Firm-Specific Changes in Labour Demand? Evidence from Idiosyncratic Export Demand Shocks," *The Review of Economic Studies*, 2023.

Hagedorn, Marcus, Tzuo Hann Law, and Iourii Manovskii, "Identifying equilibrium models of labor market sorting," *Econometrica*, 2017, *85* (1), 29–65.

Hall, Peter, "On bootstrap confidence intervals in nonparametric regression," *The Annals of Statistics*, 1992, pp. 695–711.

Hatfield, J. W. and P. R. Milgrom, "Matching with Contracts," *American Economic Review*, 2005, *95* (4), 913–935.

Hummels, David, Rasmus Jørgensen, Jakob Munch, and Chong Xiang, "The wage effects of offshoring: Evidence from Danish matched worker-firm data," *American Economic Review*, 2014, *104* (6), 1597–1629.

Huneeus, Federico, Kory Kroft, and Kevin Lim, "Earnings inequality in production networks," Technical Report, National Bureau of Economic Research 2021.

Jarosch, Gregor, Jan Sebastian Nimczik, and Isaac Sorkin, "Granular search, market structure, and wages," Technical Report, National Bureau of Economic Research 2019.

Kelso, Alexander S. Jr. and Vincent P. Crawford, "Job Matching, Coalition Formation, and Gross Substitutes," *Econometrica*, 1982, *50* (6), 1483–1504.

Klenow, Peter J and Jonathan L Willis, "Real rigidities and nominal price changes," *Economica*, 2016, *83* (331), 443–472.

Kline, Patrick, Neviana Petkova, Heidi Williams, and Owen Zidar, "Who profits from patents? rent-sharing at innovative firms," *The Quarterly Journal of Economics*, 2019, *134* (3), 1343–1404.

Kroft, Kory, Yao Luo, Magne Mogstad, and Bradley Setzler, "Imperfect competition and rents in labor and product markets: The case of the construction industry," Technical Report, National Bureau of Economic Research 2023.

Kuga, Kiyoshi and Takeshi Murota, "A Note on Definitions of Elasticity of Substitution in Many Input Case," *Metroeconomica*, 1972.

Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler, "Imperfect competition, compensating differentials, and rent sharing in the US labor market," *American Economic Review*, 2022, *112* (1), 169–212.

Lentz, Rasmus, "Sorting by search intensity," *Journal of Economic Theory*, 2010, *145* (4), 1436–1452.

Lindner, Attila, Balázs Muraközy, Balazs Reizer, and Ragnhild Schreiner, "Firm-level technological change and skill demand," 2022.

Lise, Jeremy, Costas Meghir, and Jean-Marc Robin, "Matching, sorting and wages," *Review of Economic Dynamics*, 2016, *19*, 63–87.

Maasoumi, Esfandiar and Daniel J Slottje, "Dynamics of market power and concentration profiles," *Econometric Reviews*, 2003, *22* (2), 155–177.

——— **and Henri Theil**, "The effect of the shape of the income distribution on two inequality measures," *Economics Letters*, 1979, *4* (3), 289–291.

Manning, Alan, *Monopsony in motion: Imperfect Competition in Labor Markets*, Princeton University Press, 2003.

Manski, Charles F, Daniel McFadden et al., *Structural analysis of discrete data with econometric applications*, MIT press Cambridge, MA, 1981.

McFadden, D. L., *Modeling the Choice of Residential Location*, North-Holland, 1978.

———, *Economic Models of Probabilistic Choice*, MIT Press, 1981.

Morishima, Michio, "A few suggestions on the theory of elasticity," *Keizai Hyoron (Economic Review)*, 1967, *16* (1), 144–150.

Nakamura, Emi and Dawit Zerom, "Accounting for incomplete pass-through," *The Review of Economic Studies*, 2010, *77* (3), 1192–1230.

Postel-Vinay, Fabien and Jean-Marc Robin, "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, 2002, *70* (6), 2295–2350.

Prager, Elena and Matt Schmitt, "Employer consolidation and wages: Evidence from hospitals," *American Economic Review*, 2021, *111* (2), 397–427.

Pycia, Marek, "Stability and Preference Alignment in Matching and Coalition Formation," *Econometrica*, 2012, *80* (1), 323–362.

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

Rosen, S., "The theory of equalizing differences," *Handbook of Labor Economics*, 1986, *1*, 641–692.

Roy, Andrew Donald, "Some thoughts on the distribution of earnings," *Oxford economic papers*, 1951, *3*

(2), 135–146.

Shi, Xiaoxia, Matthew Shum, and Wei Song, "Estimating semi-parametric panel multinomial choice models using cyclic monotonicity," *Econometrica*, 2018, *86* (2), 737–761.

________, ________, and ________, "Estimating Semi-Parametric Panel Multinomial Choice Models Using Cyclic Monotonicity," *Econometrica*, 2018, *86* (2), 737–761.

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

Sorkin, Isaac, "Ranking firms using revealed preference," *The Quarterly Journal of Economics*, 2018, *133* (3), 1331–1393.

Staiger, Douglas O, Joanne Spetz, and Ciaran S Phibbs, "Is there monopsony in the labor market? Evidence from a natural experiment," *Journal of Labor Economics*, 2010, *28* (2), 211–236.

Taber, Christopher and Rune Vejlin, "Estimation of a roy/search/compensating differential model of the labor market," *Econometrica*, 2020, *88* (3), 1031–1069.

Tolbert, Charles M. and Molly Sizer, "US commuting zones and labor market areas: A 1990 update," *ERS Staff Paper*, 1996.

Weyl, E Glen and Michal Fabinger, "Pass-through as an economic tool: Principles of incidence under imperfect competition," *Journal of Political Economy*, 2013, *121* (3), 528–583.

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

Appendix for
An Empirical Framework for Matching with Imperfect Competition.

Mons Chan[1], Kory Kroft[2], Elena Mattana[3], Ismael Mourifié[4]

Appendix A. Additional derivations and results.

A.1. **Optimal wage.** Under Assumption 2 (ii-a) the Karush-Kuhn-Tucker (KKT) necessary conditions for optimality of the firm's optimization problem are given by:[5]

(A-1) $\ell_{kj} + w_{kj}\frac{\partial \ell_{kj}}{\partial w_{kj}} - \lambda_j \frac{\partial \ell_{kj}}{\partial w_{kj}} F^j_k(\ell_{\cdot j}) \geq 0,$
(A-2) $w_{kj} \geq 0,$
(A-3) $w_{kj}\Big[\ell_{kj} + w_{kj}\frac{\partial \ell_{kj}}{\partial w_{kj}} - \lambda_j \frac{\partial \ell_{kj}}{\partial w_{kj}} F^j_k(\ell_{\cdot j})\Big] = 0,$
(A-4) $F^j(\ell_{\cdot j}) - Y_j \geq 0,$
(A-5) $\lambda_j \geq 0,$
(A-6) $\lambda_j\Big[F^j(\ell_{\cdot j}) - Y_j\Big] = 0$, for all $(k, j) \in (\mathcal{K} \times \mathcal{J})$.

Notice that given our ARUM and since u_{kj} is finite, $w_{kj} = 0$ implies that $\ell_{kj} = 0$. Under Assumptions 2 (i)-(ii-b), (A-4) is not violated if there exist some k such $\ell_{kj} > 0$ which means $w_{kj} > 0$ under Assumption 1. This means that each firm that is observed in this market pays a strictly positive wage to some types of worker. Let $\mathcal{C}^j \subseteq \mathcal{K}$ denote the set of worker types for whom firm j offers a strictly positive wage, $w_{kj} > 0$ which according our ARUM specification and Assumption 1 is equivalent to $s_{kj} > 0$. Then we have $\mathcal{C}^j \equiv \{k \in \mathcal{K} : s_{kj} > 0\}$. Then, (A-3) implies that (A-1) holds as an equality for all $k \in \mathcal{C}^j$ and thus $\ell_{kj} > 0$ for all $k \in \mathcal{C}^j$. We then have

$$w_{kj} = \lambda_j F^j_k(\ell_{\cdot j}) \frac{\mathcal{E}_{kj}}{1 + \mathcal{E}_{kj}}, \text{ for all } k \in \mathcal{C}^j \tag{A.1}$$

In this case, firm j optimally chooses to offer a wage equal to 0 when A-1 holds with strict inequality which corresponds to the case where the marginal cost for this type of worker exceeds the marginal product. Also, notice that all the RHS terms have to be positive to ensure that A-4 holds, which is compatible with the previous assumption used in the model.

A.2. **Recovering unobserved types.** The proposed identification strategy requires us to observe at least two time periods. We consider the following potential outcomes model:

$$Y_{it} = \sum_{j \in \mathcal{J}_0} [\ln w_{\mathbf{k}jt} + \eta_{ijt}] 1\{D_{it} = j\}, \quad t \in \{1, ..., T\} \tag{A.2}$$

[1]Department of Economics, Queen's University. 317 Dunning Hall 94 University Ave Kingston, ON, K7L 3N6, mons.chan@queensu.ca.
[2]Department of Economics, University of Toronto, & NBER. 150 St. George Street, Toronto ON M5S 3G7, Canada, kory.kroft@utoronto.ca.
[3]Department of Economics and Business Economics Aarhus University, Fuglesangs Allé 4 8210 Aarhus V Denmark, emattana@econ.au.dk.
[4]Department of Economics, Washington University in St. Louis & NBER, ismaelm@wustl.edu.
[5]Notice that in the case where the production functions are non-differentiable (for instance the Leontief Production function) sub-differential versions of KKT conditions are available and can be applied.

where Y_{it} denotes the observed log earnings of individual i at time t, and $1\{\cdot\}$ denotes the indicator function. $Y_{ijt} \equiv \ln w_{\mathbf{k}jt} + \eta_{ijt}$ denotes potential log earnings if individual i was externally assigned to work at firm j in period t. The potential outcomes are decomposed into two parts (i) $\ln w_{\mathbf{k}jt}$ is the log equilibrium wage, and (ii) η_{ijt} is measurement error or an i.i.d. worker-firm match effect realized after potential mobility across periods.

While in the main text we assumed that the worker's type k is observed by both firms and the econometrician, in general, we could allow k to consist of two subgroups of types, i.e. $k \equiv (\bar{k}, \tilde{k})$, where $\bar{k}$ is defined based on the underlying vector of characteristics $\overline{X}$ that are observed both by the econometrician and firms while $\tilde{k}$ is defined based on the set of characteristics $\widetilde{X}$ that are observable only to firms but not to the econometrician.

Let m_{it} denote the mobility variable, more precisely $m_{it} = 1$ iff $D_{it} \neq D_{it+1}$, i.e. $m_{it} = 1\{D_{it} \neq D_{it+1}\}$. Using shorthand notation $\bar{\mathbf{k}}^{\mathbf{t+1}} = (\bar{\mathbf{k}}_{\mathbf{t}}, \bar{\mathbf{k}}_{\mathbf{t+1}})$, consider the following assumption:

Assumption 4 (Time invariance, Mobility, and Serial Dependence)**.** *We impose the following restrictions.*

(i) *Time invariance of unobserved types:* $\tilde{\mathbf{k}}_{\mathbf{t}} = \tilde{\mathbf{k}}$ *for* $t \in \{1, ..., T\}$.
(ii) *Classical errors:* $(\eta_{ijt}, \eta_{ilt+1}) \perp (D_{it}, D_{it+1}) | \tilde{\mathbf{k}}, \bar{\mathbf{k}}_{\mathbf{t}}, \bar{\mathbf{k}}_{\mathbf{t+1}}$
(iii) *No serial dependence in the errors:* $\eta_{ijt} \perp \eta_{ilt+1} | \tilde{\mathbf{k}}, \bar{\mathbf{k}}_{\mathbf{t}}, \bar{\mathbf{k}}_{\mathbf{t+1}}$ *and* $\eta_{ijt} \perp \bar{\mathbf{k}}_{\mathbf{t+1}} | \tilde{\mathbf{k}}, \bar{\mathbf{k}}_{\mathbf{t}}$

Assumption 4(i) requires the unobserved types to be time invariant. In the same spirit as Burdett and Mortensen (1998) and Hagedorn et al. (2017), Assumption 4(ii) requires the errors to not be correlated with sorting and mobility decisions. The intuition is that these errors are realized after the matches between workers and firms have been formed. Assumption 4(iii) requires the measurement errors associated to a specific mover to not be serially dependent.

Under Assumption 4 we can show that

$$\begin{aligned}
&\mathbb{P}(Y_{it} \leq y_t, Y_{i,t+1} \leq y_{t+1} | D_{it} = j, D_{it+1} = l, m_{it} = 1, \bar{\mathbf{k}}^{\mathbf{t+1}} = \bar{k}^{t+1}) \\
&\qquad = \sum_{\tilde{k}} \mathbb{P}_{\tilde{k}j}(y_t | \bar{k}_t) \mathbb{P}^m_{\tilde{k}l}(y_{t+1} | \bar{k}^{t+1}) \mathbb{P}(\tilde{\mathbf{k}} = \tilde{k} | D_{it} = j, D_{it+1} = l, m_{it} = 1, \bar{\mathbf{k}}^{\mathbf{t+1}} = \bar{k}^{t+1}) \qquad \text{(A.3)}
\end{aligned}$$

where

$$\mathbb{P}_{\tilde{k}j}(y_t | \bar{k}_t) \equiv \mathbb{P}(Y_{it} \leq y_t | D_{it} = j, \tilde{\mathbf{k}} = \tilde{k}, \bar{\mathbf{k}}_{\mathbf{t}} = \bar{k}_t), \qquad \text{(A.4)}$$

$$\mathbb{P}^m_{\tilde{k}l}(y_{t+1} | \bar{k}^{t+1}) \equiv \mathbb{P}(Y_{i,t+1} \leq y_{t+1} | D_{it+1} = l, m_{it} = 1, \tilde{\mathbf{k}} = \tilde{k}, \bar{\mathbf{k}}^{\mathbf{t+1}} = \bar{k}^{t+1}). \qquad \text{(A.5)}$$

Whenever the above decomposition holds and the following three requirements hold: (i) Any two firms j and l belong to a connecting cycle as formally defined in Bonhomme et al. (2019), Definition 1, (ii) there exists some asymmetry in the worker type composition between different firms, i.e, Bonhomme et al. (2019), Assumption 3(i), and (iii) the matrix defined by the joint log earning distribution $\mathbb{P}(Y_{it} \leq y_t, Y_{i,t+1} \leq y_{t+1} | D_{it} = j, D_{it+1} = l, m_{it} = 1, \bar{\mathbf{k}}^{\mathbf{t+1}} = \bar{k}^{t+1})$ for different values of (y_t, y_{t+1}) respects a certain rank condition, i.e, Bonhomme et al. (2019), Assumption 3(ii). Then Theorem 1 of Bonhomme et al. (2019) applies and the following quantities are point identified: $\mathbb{P}_{\tilde{k}j}(y_t | \bar{k}_t)$, $\mathbb{P}^m_{\tilde{k}l}(y_{t+1} | \bar{k}_{t+1})$, and $\mathbb{P}_{jt}(\tilde{k} | \bar{k}_t) \equiv \mathbb{P}(\tilde{\mathbf{k}} = \tilde{k} | D_{it} = j, \bar{\mathbf{k}}_{\mathbf{t}} = \bar{k}_t)$.

These distributions can be parametrically estimated using the EM algorithm entertained in Bonhomme et al. (2019). Using this identification result, it is possible to recover equilibrium wages and shares that were initially unobserved to the econometrician. More precisely, we have the following result:

Proposition 3 (Identification of equilibrium wages and shares)**.** *Consider Assumption 4 holds, and the cdf of classical errors $F_{\eta_{ijt}|\mathbf{k_t}=k_t}(.)$, and $F_{\eta_{ilt+1}|\mathbf{k^{t+1}}=k^{t+1}}(.)$ are known and strictly increasing on $\mathbb{R}$. If the following quantities are point identified $\mathbb{P}_{\tilde{k}j}(y_t|\bar{k}_t)$, $\mathbb{P}^m_{\tilde{k}l}(y_{t+1}|\bar{k}_{t+1})$, $\mathbb{P}_{jt}(\tilde{k}|\bar{k}_t)$; then we have the following identification result:*

$$w_{kjt} = \exp\left\{y_t - F^{-1}_{\eta_{ijt}|\mathbf{k_t}=k_t}\left(\mathbb{P}_{\tilde{k}j}(y_t|\bar{k}_t)\right)\right\}, \tag{A.6}$$

$$w_{klt+1} = \exp\left\{y_{t+1} - F^{-1}_{\eta_{ilt+1}|\mathbf{k^{t+1}}=k^{t+1}}\left(\mathbb{P}^m_{\tilde{k}l}(y_{t+1}|\bar{k}_{t+1})\right)\right\}, \tag{A.7}$$

$$s_{kjt} = \mathbb{P}_{jt}(\tilde{k}|\bar{k}_t)\frac{s_{\bar{k}jt}}{\sum_{\mathcal{J}_0}\mathbb{P}_{jt}(\tilde{k}|\bar{k}_t)s_{\bar{k}jt}}. \tag{A.8}$$

where $s_{kjt} = \mathbb{P}(D_{it} = j|\mathbf{k_t} = k_t)$ and $s_{\bar{k}jt} = \mathbb{P}(D_{it} = j|\bar{\mathbf{k}}_\mathbf{t} = \bar{k}_t)$

Proof of Proposition 3.

$$\begin{aligned}
&\mathbb{P}(Y_{it} \le y_t, Y_{i,t+1} \le y_{t+1}|D_{it} = j, D_{it+1} = l, m_{it} = 1, \bar{\mathbf{k}}^{\mathbf{t+1}} = \bar{k}^{t+1})\\
&\quad= \sum_{\tilde{k}} \mathbb{P}(Y_{it} \le y_t, Y_{i,t+1} \le y_{t+1}|D_{it} = j, D_{it+1} = l, \tilde{\mathbf{k}} = \tilde{k}, \bar{\mathbf{k}}^{\mathbf{t+1}} = \bar{k}^{t+1})\times\\
&\qquad \overbrace{\mathbb{P}(\tilde{\mathbf{k}} = \tilde{k}|D_{it} = j, D_{it+1} = l, \bar{\mathbf{k}}_\mathbf{t} = \bar{k}_t, \bar{\mathbf{k}}_\mathbf{t+1} = \bar{k}_{t+1})}^{P(\tilde{k}|j,l,\bar{k}^{t+1})}\\
&= \sum_{\tilde{k}} \mathbb{P}(\ln w_{\mathbf{k}jt} + \eta_{ijt} \le y_t, \ln w_{\mathbf{k}j,t+1} + \eta_{ilt+1} \le y_{t+1}|D_{it} = j, D_{it+1} = l, \tilde{\mathbf{k}} = \tilde{k}, \bar{\mathbf{k}}^{\mathbf{t+1}} = \bar{k}^{t+1}) \times P(\tilde{k}|j,l,\bar{k}^{t+1})\\
&\quad= \sum_{\tilde{k}} \mathbb{P}\left(\ln w_{\mathbf{k}jt} + \eta_{ijt} \le y_t, \ln w_{\mathbf{k}j,t+1} + \eta_{ilt+1} \le y_{t+1}|\tilde{\mathbf{k}} = \tilde{k}, \bar{\mathbf{k}}^{\mathbf{t+1}} = \bar{k}^{t+1}\right) \times P(\tilde{k}|j,l,\bar{k}^{t+1})\\
&= \sum_{\tilde{k}} \mathbb{P}\left(\ln w_{\mathbf{k}jt} + \eta_{ijt} \le y_t|\tilde{\mathbf{k}} = \tilde{k}, \bar{\mathbf{k}}^{\mathbf{t+1}} = \bar{k}^{t+1}\right)\times\mathbb{P}\left(\ln w_{\mathbf{k}j,t+1} + \eta_{ilt+1} \le y_{t+1}|\tilde{\mathbf{k}} = \tilde{k}, \bar{\mathbf{k}}^{\mathbf{t+1}} = \bar{k}^{t+1}\right)\times P(\tilde{k}|j,l,\bar{k}^{t+1})\\
&\quad= \sum_{\tilde{k}} \mathbb{P}\left(\ln w_{\mathbf{k}jt} + \eta_{ijt} \le y_t, \ln w_{\mathbf{k}j,t+1} + \eta_{ilt+1} \le y_{t+1}|\tilde{\mathbf{k}} = \tilde{k}, \bar{\mathbf{k}}^{\mathbf{t+1}} = \bar{k}^{t+1}\right) \times P(\tilde{k}|j,l,\bar{k}^{t+1})\\
&= \sum_{\tilde{k}} \mathbb{P}\left(Y_{it} \le y_t|D_{it} = j, \tilde{\mathbf{k}} = \tilde{k}, \bar{\mathbf{k}}_\mathbf{t} = \bar{k}_t\right)\times\mathbb{P}\left(Y_{i,t+1} \le y_{t+1}|D_{it+1} = l, m_{it} = 1, \tilde{\mathbf{k}} = \tilde{k}, \bar{\mathbf{k}}^{\mathbf{t+1}} = \bar{k}^{t+1}\right)\times P(\tilde{k}|j,l,\bar{k}^{t+1})
\end{aligned}$$

Now, we have

$$\begin{aligned}
&\mathbb{P}_{\tilde{k}j}(y_t|\bar{k}_t) \equiv \mathbb{P}(Y_{it} \le y_t|D_{it} = j, \tilde{\mathbf{k}} = \tilde{k}, \bar{\mathbf{k}}_\mathbf{t} = \bar{k}_t)\\
&= \mathbb{P}(\ln w_{\mathbf{k}jt} + \eta_{ijt} \le y_t|D_{it} = j, \tilde{\mathbf{k}} = \tilde{k}, \bar{\mathbf{k}}_\mathbf{t} = \bar{k}_t) = \mathbb{P}(\ln w_{\mathbf{k}jt} + \eta_{ijt} \le y_t|\tilde{\mathbf{k}} = \tilde{k}, \bar{\mathbf{k}}_\mathbf{t} = \bar{k}_t) = \mathbb{P}(\eta_{ijt} \le y_t - \ln w_{\mathbf{k}jt}|\tilde{\mathbf{k}} = \tilde{k}, \bar{\mathbf{k}}_\mathbf{t} = \bar{k}_t)\\
&\qquad = F_{\eta_{ijt}|\bar{\mathbf{k}}_\mathbf{t}=\bar{k}_t}(y_t - \ln w_{\mathbf{k}jt})
\end{aligned}$$

We can then easily recover the first result by inverting the last equation and obtain: $w_{kjt} = \exp\left\{y_t - F^{-1}_{\eta_{ijt}|\bar{\mathbf{k}}_\mathbf{t}=\bar{k}_t}\left(\mathbb{P}_{\tilde{k}j}(y_t|\bar{k}_t)\right)\right\}$. The second equality of the proposition could be derived analogously. For the last equality we have:

$$\begin{aligned}\mathbb{P}(D_{it}=j|\tilde{\mathbf{k}}=\tilde{k},\bar{\mathbf{k}}_{\mathbf{t}}=\bar{k}_t) &= \frac{\mathbb{P}(\tilde{\mathbf{k}}=\tilde{k}|D_{it}=j,\bar{\mathbf{k}}_{\mathbf{t}}=\bar{k}_t)\times\mathbb{P}(D_{it}=j|\tilde{\mathbf{k}}=\tilde{k},\bar{\mathbf{k}}_{\mathbf{t}}=\bar{k}_t)}{\mathbb{P}(\tilde{\mathbf{k}}=\tilde{k}|\bar{\mathbf{k}}_{\mathbf{t}}=\bar{k}_t)} \\ &= \frac{\mathbb{P}(\tilde{\mathbf{k}}=\tilde{k}|D_{it}=j,\bar{\mathbf{k}}_{\mathbf{t}}=\bar{k}_t)\times\mathbb{P}(D_{it}=j|\tilde{\mathbf{k}}=\tilde{k},\bar{\mathbf{k}}_{\mathbf{t}}=\bar{k}_t)}{\sum_j\mathbb{P}(\tilde{\mathbf{k}}=\tilde{k}|D_{it}=j,\bar{\mathbf{k}}_{\mathbf{t}}=\bar{k}_t)\times\mathbb{P}(D_{it}=j|\tilde{\mathbf{k}}=\tilde{k},\bar{\mathbf{k}}_{\mathbf{t}}=\bar{k}_t)}\end{aligned}$$

□

Parametric estimation and EM algorithm. For practical purposes, we impose a normality distribution for the classical errors, then $\ln w_{kjt}+\eta_{ijt}|\mathbf{k^t}=k^t\sim N(\ln w_{kjt},\varrho_{kjt})$ and $\ln w_{klt}+\eta_{ilt+1}|\mathbf{k^{t+1}}=k^{t+1}\sim N(\ln w_{klt+1},\varrho_{klt+1})$. Let $\tilde{K}$ denote the number of unobserved types, $C_{\bar{k}^t}$ be a set of firms that have been hiring workers of observable types $\bar{k}^t$ over the two periods t and $t+1$ and belonging to a connecting cycle as defined in Bonhomme et al. (2019). $N^m_{\bar{k}^t}$ denotes the number of movers with observable types $\bar{k}^t$. First, we consider the following log-likelihood function for job movers:

$$\sum_{i=1}^{N^m_{\bar{k}^t}}\sum_{j\in C_{\bar{k}^t}}\sum_{l\in C_{\bar{k}^t}}\ln\left(\sum_{\tilde{k}=1}^{\tilde{K}}p_{\tilde{k}jl}\frac{1}{\sqrt{4\pi^2\varrho_{(\tilde{k},\bar{k}_t)jt}\varrho_{(\tilde{k},\bar{k}_t)lt+1}}}e^{-\frac{\left(y_{it}-\ln w_{(\tilde{k},\bar{k}_t)jt}\right)^2}{2\varrho^2_{(\tilde{k},\bar{k}_t)jt}}-\frac{\left(y_{it+1}-\ln w_{(\tilde{k},\bar{k}_t)lt+1}\right)^2}{2\varrho^2_{(\tilde{k},\bar{k}_t)lt+1}}}\right) \tag{A.9}$$

where $\hat{w}_{(\tilde{k},\bar{k}_t)jt},\hat{w}_{(\tilde{k},\bar{k}_t)lt+1},\hat{\varrho}_{(\tilde{k},\bar{k}_t)jt},\hat{\varrho}_{(\tilde{k},\bar{k}_t)lt+1}$, and $\hat{p}_{\tilde{k}jl}$ for $\tilde{k}=1,...,\tilde{K}$ are estimated by maximizing (*A.10*) using the EM algorithm.

Second, we consider the log-likelihood of the for all workers at the period t:

$$\sum_{i=1}^{N_{\bar{k}^t}}\sum_{j\in C_{\bar{k}_t}}\ln\left(\sum_{\tilde{k}=1}^{\tilde{K}}q_{\tilde{k}jt}\frac{1}{\sqrt{4\pi^2\hat{\varrho}_{(\tilde{k},\bar{k}_t)jt}}}e^{-\frac{\left(y_{it}-\ln\hat{w}_{(\tilde{k},\bar{k}_t)jt}\right)^2}{2\hat{\varrho}^2_{(\tilde{k},\bar{k}_t)jt}}}\right) \tag{A.10}$$

where $N_{\bar{k}_t}$ denotes the number of workers with observable types $\bar{k}_t$, and $q_{\tilde{k}jt}\equiv\mathbb{P}_{jt}(\tilde{k}|\bar{k}_t)$. Again we estimate $\hat{q}_{\tilde{k}jt}$ by maximizing eq (A.10) using the EM algorithm. Then we use eq (A.8) to recover $\hat{s}_{kjt}$.

Given employment shares s_{kjt} for each firm and worker type, we can then obtain the total quantity of each worker type in the population, $m_{kt}=\sum_j\ell_{kjt}$, as the (year-by-year) solution to an overdetermined system of linear equations: $S_tm_t=\mu_t$. Here S_t is the known $J\times K$ matrix of worker type shares s_{kjt} at each firm in period t, μ_t is the known $J\times 1$ vector of total employment $\mu_{jt}=\sum_{k\in\mathcal{C}^j_t}\ell_{kjt}$ at each firm, and m_t is the unknown $K\times 1$ vector of individuals m_{kt} of each type k. If both S_t and the associated augmented matrix have rank equal to K, then there will be a unique solution which provides m_{kt} for each period t[6]. We can then obtain $\ell_{kjt}=s_{kjt}m_{kt}$ for each firm, type and year.

Given that we have recovered the equilibrium wages and shares, and number of matches, these objects can then be used to recover the model parameters.

A.3. **Identifying the Labor Supply Parameters.**

[6]This is the Rouché-Capelli theorem.

A.3.1. *Estimating the Supply Equation.* The baseline labor supply equation from the model is

$$\ln \frac{s_{kjt}}{s_{k0t}} = \overline{u}_k + \beta_{1k} \ln \frac{w_{kjt}}{w_{k0t}} + \sum_{g=1}^{G} \tilde{\sigma}_{kg} \ln s_{kj|gt} \mathbb{1}_{j|g} + \ln u_{kjt} \tag{A.11}$$

where $\tilde{\sigma}_{kg} \equiv (1 - 1/\sigma_{kg})$. Define $\mathbb{1}_{j|g} = 1$ if $j \in g$ and 0 else.

The identification challenge is that both the wage and inside share are potentially correlated with the unobserved amenities and thus endogenous. To address this challenge, we propose an instrumental variables (IV) strategy which leverages exogenous variation in firm productivity. Before discussing this IV strategy, we review candidate instruments which we considered.

One source of instruments relies on strategic interactions between firms in wage setting. In the presence of strategic interactions, the number and characteristics of other firms in a given labor market can be used as instruments. These so-called "BLP instruments" are very common in the industrial organization literature in the context of the product market where the characteristics and number of competing products are used as instruments for prices (see Berry et al. (1995) (BLP) for the canonical example). In a labor market context, possible BLP instruments might include the number of firms, average size, or average value-added per worker of other firms in the labor market. Azar et al. (2022a) use the number of vacancies and log employment of competing firms as instruments for advertised wages on a job posting website. In results not reported, we consider the available BLP instruments in our data, such as the number of firms in the same market, and found that they were not sufficiently strong. Thus, we do not emphasize BLP instruments in our setting.

A second source of wage instruments exploits "uniform wage setting" whereby firms set wages similarly across local labor markets (Hazell et al., 2022). As suggested by Azar et al. (2022a), this implies that the wage a firm pays in a given market may be driven by the labor market conditions that same firm faces in other markets. We thus considered Hausman instruments for w_{kjg} in market g using the average predicted wage across all markets that firm operates in other than g[7]. In results not reported, we implemented this approach, following Azar et al. (2022a), but generally found that these instruments were too weak in our setting.

Finally, we considered a shift-share IV approach following Hummels et al. (2014) and Garin and Silvério (2023) to estimate labor supply. To construct this instrument, we rely on firm-product-country level yearly foreign trade data from Statistics Denmark register UHDI and bilateral trade flows from the BACI dataset. We find that our labor supply parameters are comparable to our main estimates reported in Table D.5. We do not emphasize these estimates as much in the paper since we are only able to construct the instrument for the small share of the firms in our sample who export. These results are available upon request.

For any of those approaches, let's present how the parameters can be consistently estimated.

A.3.2. *Estimating the Supply Equation in Changes.* We can rewrite the supply equation in changes as

$$\Delta_{e,e'} \ln \frac{s_{kj|gt}}{s_{k0t}} = \beta_{0k} + \beta_{1k} \Delta_{e,e'} \ln \frac{w_{kj|gt}}{w_{k0t}} + \sum_{g=1}^{G} \tilde{\sigma}_{kgt} \Delta_{e,e'} \ln s_{kj|gt} \mathbb{1}_{j|g} + \Delta_{e,e'} \ln u_{kj|gt} \tag{A.12}$$

where $\Delta_{e,e'} x_t \equiv x_{t+e} - x_{t-e'}$.

[7] We also exclude markets in the same municipality or industry as g.

For ease of notation, we will fix a labor type k (dropping the notation) and pool observations across firms and years (and markets), replacing indices (j,t) with a single index $n \in 1, ..., N$ representing total number of observations for labor type k. We define $\tilde{s}_n = \ln \frac{s_{jt}}{s_{0t}}$, $\tilde{w}_n = \ln \frac{w_{jt}}{w_{0t}}$, $\tilde{i}_{ng} = \ln s_{j|gt}\mathbb{1}_{j|g}$, and $\tilde{u}_n = \ln u_{jt}$. We can write this equation in matrix notation as

$$\underset{N\times 1}{\mathbf{S^\Delta}} = \underset{N\times 1}{\mathbf{X_0}}\beta_0 + \underset{N\times(G+1)}{\mathbf{X_1^\Delta}}\underset{(G+1)\times 1}{\boldsymbol{\beta}} + \underset{N\times 1}{\mathbf{U^\Delta}} \tag{A.13}$$

where $\mathbf{X_0}$ is a column vector of 1's,

$$\mathbf{S^\Delta} = \begin{pmatrix} \Delta_{e,e'}\tilde{s}_1 \\ \Delta_{e,e'}\tilde{s}_2 \\ \vdots \\ \Delta_{e,e'}\tilde{s}_N \end{pmatrix}, \quad \mathbf{X_1^\Delta} = \begin{pmatrix} \Delta_{e,e'}\tilde{w}_1 & \Delta_{e,e'}\tilde{i}_{11} & \cdots & \Delta_{e,e'}\tilde{i}_{1G} \\ \Delta_{e,e'}\tilde{w}_2 & \Delta_{e,e'}\tilde{i}_{21} & \cdots & \Delta_{e,e'}\tilde{i}_{2G} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_{e,e'}\tilde{w}_N & \Delta_{e,e'}\tilde{i}_{N1} & \cdots & \Delta_{e,e'}\tilde{i}_{NG} \end{pmatrix}, \quad \mathbf{U^\Delta} = \begin{pmatrix} \Delta_{e,e'}\tilde{u}_1 \\ \Delta_{e,e'}\tilde{u}_2 \\ \vdots \\ \Delta_{e,e'}\tilde{u}_N \end{pmatrix}$$

Define $(\mathbf{W}^\Delta)^T = (\Delta_{e,e'}\tilde{w}_1, ..., \Delta_{e,e'}\tilde{w}_N)$ and $(\mathbf{I_g}^\Delta)^T = (\Delta_{e,e'}\tilde{i}_1 g, ..., \Delta_{e,e'}\tilde{i}_N g)$. Suppose we now want to use variable Δr_n to instrument for $\Delta_{e,e'}\tilde{w}_n$, and variable Δf_{ng} to instrument for $\Delta_{e,e'}\tilde{i}_{ng}$. Here, Δr_n is the one-period change in (log) firm revenues and Δf_{ng} is the one-period change in the (log) inside share in market g, where as above n indexes across j and t. Define the matrix of instruments $\mathbf{Z^\Delta}$ as

$$\mathbf{Z^\Delta} = \begin{pmatrix} \mathbf{R^\Delta} & \mathbf{F_1^\Delta} & \cdots & \mathbf{F_G^\Delta} \end{pmatrix} = \begin{pmatrix} \Delta r_1 & \Delta f_{11} & \cdots & \Delta f_{1G} \\ \Delta r_2 & \Delta f_{21} & \cdots & \Delta f_{2G} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta r_N & \Delta f_{N1} & \cdots & \Delta f_{NG} \end{pmatrix}$$

Given the intercept term, as above, we can write the instrumental variable estimator for $\boldsymbol{\beta}$ with the equation in changes as

$$\widehat{\boldsymbol{\beta}}^\Delta = \mathrm{Cov}(\mathbf{Z^\Delta}, \mathbf{X_1^\Delta})^{-1}\mathrm{Cov}(\mathbf{Z^\Delta}, \mathbf{S^\Delta}) \tag{A.14}$$

$$= \begin{pmatrix} \mathbf{C^\Delta_{RW}} & \mathbf{C^\Delta_{RI}} \\ \mathbf{C^\Delta_{FW}} & \mathbf{C^\Delta_{FI}} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{C^\Delta_{RS}} \\ \mathbf{C^\Delta_{FS}} \end{pmatrix} \tag{A.15}$$

where

$$\underset{1\times 1}{\mathbf{C^\Delta_{RW}}} = \mathrm{Cov}(\mathbf{R^\Delta}, \mathbf{W^\Delta}), \quad \underset{1\times G}{\mathbf{C^\Delta_{RI}}} = \Big(\mathrm{Cov}(\mathbf{R^\Delta}, \mathbf{I_1^\Delta}) \cdots \mathrm{Cov}(\mathbf{R^\Delta}, \mathbf{I_G^\Delta})\Big) \tag{A.16}$$

and

$$\underset{G\times 1}{\mathbf{C^\Delta_{FW}}} = \begin{pmatrix} \mathrm{Cov}(\mathbf{F_1^\Delta}, \mathbf{W^\Delta}) \\ \vdots \\ \mathrm{Cov}(\mathbf{F_G^\Delta}, \mathbf{W^\Delta}) \end{pmatrix}, \quad \underset{G\times G}{\mathbf{C^\Delta_{FI}}} = \begin{pmatrix} \mathrm{Cov}(\mathbf{F_1^\Delta}, \mathbf{I_1^\Delta}) \cdots \mathrm{Cov}(\mathbf{F_1^\Delta}, \mathbf{I_G^\Delta}) \\ \vdots \ddots \vdots \\ \mathrm{Cov}(\mathbf{F_G^\Delta}, \mathbf{I_1^\Delta}) \cdots \mathrm{Cov}(\mathbf{F_G^\Delta}, \mathbf{I_G^\Delta}) \end{pmatrix} \tag{A.17}$$

and finally

$$\underset{1\times 1}{\mathbf{C^\Delta_{RS}}} = \mathrm{Cov}(\mathbf{R^\Delta}, \mathbf{S^\Delta}), \quad \underset{G\times 1}{(\mathbf{C^\Delta_{FS}})^T} = \Big(\mathrm{Cov}(\mathbf{F_1^\Delta}, \mathbf{S^\Delta}) \cdots \mathrm{Cov}(\mathbf{F_G^\Delta}, \mathbf{S^\Delta})\Big) \tag{A.18}$$

What comes next requires a few assumptions:

Assumption 5. *The instruments are predetermined. i.e.:* $C^\Delta_{RU} \equiv \mathit{Cov}(\mathbf{R^\Delta}, \mathbf{U^\Delta}) = 0$ *and* $C^\Delta_{FU} \equiv \mathit{Cov}(\mathbf{F^\Delta}, \mathbf{U^\Delta}) = \mathbf{0}$.

Assumption 6. *The instruments are valid and correlated with the endogenous regressors. i.e.: The* $G \times G$ *matrix* $\mathbb{E}[(\mathbf{Z^\Delta})'\mathbf{X_1^\Delta}]$ *is full column rank.*

These two assumptions are similar to assumptions made in Lamadon et al. (2022) and Kroft et al. (2023), who also estimate labor supply systems in changes. Specifically, assumptions 5 and 6 together encompass assumption 3 in Kroft et al. (2023). Assumptions 5 and 6 are satisfied for each instrument z_{jkt} if (briefly using full notation) $\exists\, e\,,\, e' > 0$ such that $\mathrm{Cov}(\tilde{\gamma}_{kjt+e} - \tilde{\gamma}_{kjt-e'}, \Delta z_{jkt}) \neq 0$ and $\mathrm{Cov}(\ln u_{kjt+e} - \ln u_{kjt-e'}, \Delta z_{jkt}) = 0$. The first is satisfied if the firm productivity process is sufficiently persistent (i.e.: δ is sufficiently close to 1 under the AR(1) assumptions in section 5.2). The second is satisfied if the amenity process is sufficiently transitory. Lamadon et al. (2022) and Kroft et al. (2023) argue that unobserved firm-specific job amenity shocks are well approximated by an MA(1) process. Under this specification, $e \geq 2$ and $e' \geq 3$ satisfy the exclusion restrictions.

Given these assumptions, the estimator becomes

$$\widehat{\boldsymbol{\beta}}^{\Delta} = \begin{pmatrix} \widehat{\beta_1}^{\Delta} \\ \widehat{\tilde{\sigma}}^{\Delta} \end{pmatrix} = \begin{pmatrix} \frac{1}{\overline{C}^{\Delta}} \left(\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{RS}} - \mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{RI}}(\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FI}})^{-1}\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FS}}\right) \\ \frac{1}{\overline{C}^{\Delta}} \left((\overline{\mathbf{C}}^{\boldsymbol{\Delta}}(\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FI}})^{-\mathbf{1}} + (\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FI}})^{-\mathbf{1}}\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FW}}\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{RI}}(\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FI}})^{-1})\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FS}} - (\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FI}})^{-1}\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FW}}\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{RS}}\right) \end{pmatrix} \tag{A.19}$$

where $\overline{C}^{\Delta} \equiv \mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{RW}} - \mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{RI}}(\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FI}})^{-\mathbf{1}}\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FW}}$ is a non-zero scalar, since assumption 6 implies that $\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{RW}}$ is non-zero and $\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FI}}$ is invertible. We can then state the following result:

Proposition 4. *Under Assumptions 5 and 6, $\widehat{\beta}^{\Delta}$ recovers β.*

Proof. By equation A.13 we have:

$$\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{RS}} = \mathrm{Cov}(\mathbf{R}^{\boldsymbol{\Delta}}, \mathbf{S}^{\boldsymbol{\Delta}}) = \mathrm{Cov}(\mathbf{R}^{\boldsymbol{\Delta}}, \mathbf{W}^{\boldsymbol{\Delta}}\beta_1^{\Delta} + \mathbf{I}\tilde{\boldsymbol{\sigma}}^{\boldsymbol{\Delta}} + \mathbf{U}^{\boldsymbol{\Delta}})$$

and

$$\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FS}} = \mathrm{Cov}(\mathbf{F}^{\boldsymbol{\Delta}}, \mathbf{S}^{\boldsymbol{\Delta}}) = \mathrm{Cov}(\mathbf{F}^{\boldsymbol{\Delta}}, \mathbf{W}^{\boldsymbol{\Delta}}\beta_1^{\Delta} + \mathbf{I}\tilde{\boldsymbol{\sigma}}^{\boldsymbol{\Delta}} + \mathbf{U}^{\boldsymbol{\Delta}})$$

By equation A.19 and assumption 6, the estimator $\widehat{\beta_1}^{\Delta}$ is thus

$$\begin{aligned}
\widehat{\beta_1}^{\Delta} &= \frac{1}{\overline{C}^{\Delta}} \left(\beta^{\boldsymbol{\Delta}}_{\mathbf{1}}\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{RW}} + \mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{RI}}\tilde{\boldsymbol{\sigma}}^{\boldsymbol{\Delta}} + \mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{RU}} - \mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{RI}}(\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FI}})^{-\mathbf{1}}(\beta^{\boldsymbol{\Delta}}_{\mathbf{1}}\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FW}} + \mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FI}}\tilde{\boldsymbol{\sigma}}^{\boldsymbol{\Delta}} + \mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FU}})\right) \\
&= \frac{1}{\overline{C}^{\Delta}} \left(\beta^{\boldsymbol{\Delta}}_{\mathbf{1}}(\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{RW}} - \mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{RI}}(\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FI}})^{-\mathbf{1}}\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FW}}) + (\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{RI}} - \mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{RI}}(\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FI}})^{-\mathbf{1}}\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FI}})\tilde{\boldsymbol{\sigma}}^{\boldsymbol{\Delta}}\right) \\
&= \beta_1^{\Delta}\frac{\overline{C}^{\Delta}}{\overline{C}^{\Delta}} + \frac{1}{\overline{C}^{\Delta}}\mathbf{0}\tilde{\boldsymbol{\sigma}}^{\boldsymbol{\Delta}} \\
&= \beta_1^{\Delta}
\end{aligned}$$

where the second equation is due to assumption 5, and the third equation is due to the definition of $\overline{C}^{\Delta}$. Similarly, by assumption 6, for $\tilde{\boldsymbol{\sigma}}^{\boldsymbol{\Delta}}$ we have

$$\begin{aligned}
\widehat{\tilde{\boldsymbol{\sigma}}^{\boldsymbol{\Delta}}} &= \frac{1}{\overline{C}^{\Delta}} \Big((\overline{\mathbf{C}}^{\boldsymbol{\Delta}}(\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FI}})^{-1} + (\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FI}})^{-\mathbf{1}}\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FW}}\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{RI}}(\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FI}})^{-1})(\beta^{\boldsymbol{\Delta}}_{\mathbf{1}}\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FW}} + \mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FI}}\tilde{\boldsymbol{\sigma}}^{\boldsymbol{\Delta}} + \mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FU}}) + \\
&\qquad - (\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FI}})^{-\mathbf{1}}\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FW}}(\beta^{\boldsymbol{\Delta}}_{\mathbf{1}}\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{RW}} + \mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{RI}}\tilde{\boldsymbol{\sigma}}^{\boldsymbol{\Delta}} + \mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{RU}}) \\
&= \frac{\overline{C}^{\Delta}}{\overline{C}^{\Delta}}\tilde{\boldsymbol{\sigma}}^{\boldsymbol{\Delta}} + \frac{1}{\overline{C}^{\Delta}}\beta_1^{\Delta}(\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FI}})^{-\mathbf{1}}\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FW}}(\overline{C}^{\Delta} + \mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{RI}}(\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FI}})^{-\mathbf{1}}\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FW}} - \mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{RW}}) \\
&= \tilde{\boldsymbol{\sigma}}^{\boldsymbol{\Delta}} + \frac{1}{\overline{C}^{\Delta}}\beta_1^{\Delta}(\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FI}})^{-\mathbf{1}}\mathbf{C}^{\boldsymbol{\Delta}}_{\mathbf{FW}}(\overline{C}^{\Delta} - \overline{C}^{\Delta}) \\
&= \tilde{\boldsymbol{\sigma}}^{\boldsymbol{\Delta}}
\end{aligned}$$

where again the second equality is due to assumption 5 and the third equality is due to the definition of $\overline{C}^{\Delta}$. □

A.4. **Multi-Equation GMM Approach to Estimating Production Parameters.** Estimating equation 5.6 is not straight forward. We cannot use an equation-by-equation approach as we do for the labor supply equation due to the presence of common parameters across equations. While there are only $K+1$ parameters to estimate ($\rho_k \, \forall \, k$ and δ), there are $K*(K-1)/2$ equations which could be used to estimate the parameters, with no obvious guidance on which to use. Since not all firms employ every labor type, any subset of equations will somewhat arbitrarily ignore the contribution of some firms. If all firms employed some base type of labor, all the labor ratio equations could be cast in terms of that type. However this is not the case, so an alternative is to use all $K*(K-1)/2$ equations in a multi-equation GMM estimator. Another possible approach would be to treat the multi-equation GMM system non-linearly and estimate the $K+1$ parameters directly. This would require $K+1$ instruments, for which the obvious choices are lagged labor and wages for each labor type. However, due to the size of the problem this may be intractable.

The approach we take is to treat the system as a set of linear equations with cross-equation parameter restrictions, estimating the compound parameters (such as $\delta(\rho_k-1)$) and then calculating the structural parameters post-estimation. This has the advantage of being much faster, and also allows specification testing of the model assumptions (since we can test if our estimates of $\delta(\rho_k-1)$ equal the product of our estimates of δ and (ρ_k-1)). Functionally, we estimate $K*(K-1)/2$ equations, where each equation (for all a,b in the set of labor types) takes the following form:

$$\begin{aligned} d_{kjt}d_{hjt}\log\frac{\tilde{w}_{ajt}}{\tilde{w}_{bjt}} &= \sum_k \mathbb{1}_{k=a} d_{kjt}\left[\beta_k^1 \log \ell_{kjt} - \beta_k^2 \log \mu_{kjt-1}\right] \\ &\quad - \sum_h \mathbb{1}_{h=b} d_{hjt}\left[\beta_h^1 \log \ell_{hjt} - \beta_h^2 \log \mu_{hjt-1}\right] \\ &\quad + \sum_{k,h,t} \mathbb{1}_{k=a}\mathbb{1}_{h=b} d_{kjt}d_{hjt}\left[\delta \log \frac{\tilde{w}_{kjt-1}}{\tilde{w}_{hjt-1}} + c_{kht}\right] + \eta_{abjt} \end{aligned} \tag{A.20}$$

where $\beta_k^1 \equiv (\rho_k-1)$, $\beta_k^2 \equiv \delta(\rho_k-1)$, and d_{kjt} is an indicator variable which equals 1 if firm j employs labor type k in periods t and $t-1$. This is similar to a "multivariate" regression where all the same regressors appear on the RHS of every equation. We now have $2K+1$ parameters to estimate, and thus need $2K+1$ instruments. Here we use lagged labor μ_{kjt-1}, lagged wages w_{kjt-1}, plus squares of both, giving us an overidentified system which we estimate using linear GMM (essentially 2SLS). Note that this approach allows for arbitrary cross-equation patterns of correlation between the error terms η_{abjt}.

Appendix B. Proofs of the main text results

B.1. Proof of Theorem 1. Fixed point representation of the existence of an equilibrium.

Recall that Assumptions 1, and 2, the optimal wage (eq 2.7) can be equivalently rewritten as

$$w_{kj} = \lambda_j F_k^j(\ell_{\cdot j}(w)) \frac{\mathcal{E}_{kj}(w)}{1+\mathcal{E}_{kj}(w)} \equiv B_{kj}(w), \quad \forall (k,j) \in \mathcal{K} \times \mathcal{J}. \tag{B.1}$$

Let $B(w) \equiv (B_{11}(.), ..., B_{KJ}(.))$. With this representation, showing the existence of an equilibrium matching is equivalent to show that the mapping $B(w)$ admits at least a fixed point, i.e. w^{eq}, such that $B(w^{eq}) = w^{eq}$ and then $s_{kj}(w^{eq}) = \frac{\partial G_{k\cdot}(v_{k\cdot})}{\partial v_{kj}}|_{v_{kj}=v_{kj}^{eq}}$ where $v_{kj}^{eq} \equiv \beta_{kj} \ln w_{kj}^{eq} + \ln u_{kj}$.

Let $\mathbb{T}_0 = \{w : 0 \leq w_{11} \leq \bar{\lambda}\bar{F}', ..., 0 \leq w_{KJ} \leq \bar{\lambda}\bar{F}'\}$, be a closed and bounded rectangular region in $\mathbb{R}^{KJ}$.

Step 0: Let $\underline{\xi}^t = (\underline{\xi}_1^t, ..., \underline{\xi}_{I+J}^t)$ and $\overline{\xi}^t = (\overline{\xi}_1^t, ..., \overline{\xi}_{I+J}^t)$ be vectors of arbitrarily small non-negative constants such that $\underline{\xi}_{kj}^t \leq w \leq \bar{\lambda}\bar{F}' - \overline{\xi}_{kj}^t$ for all $(k,j) \in \mathcal{K} \times \mathcal{J}$. $\underline{\xi}^t$ is chosen such that some of those components are strictly positive, which is ensured by the fact that under Assumptions 1, and 2, $\mathcal{C}^j \neq \{\emptyset\}$ for each $j \in J$. And define, $\mathbb{T}_\xi^t = \{w : \underline{\xi}_{11}^t \leq w_{11} \leq \bar{\lambda}\bar{F}' - \overline{\xi}_{11}^t, ..., \underline{\xi}_{KJ}^t \leq w_{KJ} \leq \bar{\lambda}\bar{F}' - \overline{\xi}_{KJ}^t\}$. Under Assumptions 1, and 2, also given that $B_{kj}(w)$ are continuous functions on a compact set $\mathbb{T}_\xi^t$ and $\lambda_j < \bar{\lambda}$, there exist vectors of non-negative constants (some strictly positive) $\underline{\eta}^t = (\underline{\eta}_{11}^t, ..., \underline{\eta}_{KJ}^t)$ and $\overline{\eta}^t = (\overline{\eta}_{11}^t, ..., \overline{\eta}_{KJ}^t)$ such that $\underline{\eta}_{kj}^t \leq B_{kj}(w) \leq \bar{\lambda}\bar{F}' - \overline{\eta}_{kj}^t$ for all $(k,j) \in \mathcal{K} \times \mathcal{J}$. More precisely, just take $\underline{\eta}_{kj}^t = \inf_{w \in \mathbb{T}_\xi^t} B_{kj}(w)$, and $\overline{\eta}_{kj}^t = \bar{\lambda}\bar{F}' - \sup_{w \in \mathbb{T}_\xi^t} B_{kj}(w)$, for all $(k,j) \in \mathcal{K} \times \mathcal{J}$.

Step 1: Define $\underline{\xi}_i^{t+1} = \min(\underline{\xi}_i^t, \underline{\eta}_i^t)$ for for $i = 11, ..., KJ$ and $\overline{\xi}_i^{t+1} = \min(\overline{\xi}_i^t, \overline{\eta}_i^t)$ for $i = 11, ..., KJ$.

Step 2: If $\underline{\xi}_i^{t+1} = \underline{\xi}_i^t$ and $\overline{\xi}_i^{t+1} = \overline{\xi}_i^t$ then stop the iteration and define $\underline{\epsilon}_i = \underline{\xi}_i^{t+1}$, $\bar{\epsilon}_i = \overline{\xi}_i^{t+1}$.

Step 3: If $\underline{\xi}_i^{t+1} \neq \underline{\xi}_i^t$ or $\overline{\xi}_i^{t+1} \neq \overline{\xi}_i^t$ then $t \leftarrow t+1$ and go back to step 0.

By construction $\underline{\xi}_i^t$ and $\overline{\xi}_i^t$ are decreasing positive sequences bounded from below by 0 then converge. So, when the iteration will stop in **Step 2**, let $\mathbb{T}_\epsilon = \{w : \underline{\epsilon}_{11} \leq w_{11} \leq \bar{\lambda}\bar{F}' - \bar{\epsilon}_{11}, ..., \underline{\epsilon}_{KJ} \leq w_{KJ} \leq \bar{\lambda}\bar{F}' - \bar{\epsilon}_{KJ}\}$ be a closed and bounded rectangular region in $\mathbb{R}^{KJ}$.

$B(w)$ is a continuously differentiable mapping such that $B(w)$: $\mathbb{T}_\epsilon \to \mathbb{T}_\epsilon$. Thus, the existence of a wage equilibrium w^{eq} exists by invoking the Brouwer fixed-point theorem. And then by construction we have the existence of (s^{eq}, w^{eq}).

B.2. Proof of Theorem 2. Let's define

$$\delta_{kj}(w) \equiv w_{kj} - \lambda_j F_k^j(\ell_{\cdot j}(w)) \frac{\mathcal{E}_{kj}(w)}{1+\mathcal{E}_{kj}(w)}, \quad \forall (k,j) \in \mathcal{K} \times \mathcal{J}. \tag{B.2}$$

$\delta(w) = (\delta_{11}(w), ..., \delta_{KJ}(w)) : \mathbb{T}_\epsilon \subseteq \mathbb{R}^{KJ} \longrightarrow \mathbb{R}^{KJ}$. Theorem 1 shows that an equilibrium is exist, showing the uniqueness is equivalent to show the global univalence of the mapping $\delta(w)$. Under Assumptions 1, and 2, $\delta(w)$ is continuously differentiable. Let $\mathbb{J}_\delta(w)$ be its Jacobian matrix, $\underset{(KJ \times KJ)}{\mathbb{J}_\delta(w)} = \begin{pmatrix} \frac{\partial \delta_{11}}{\partial w_{11}} & \cdots & \frac{\partial \delta_{11}}{\partial w_{KJ}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \delta_{KJ}}{\partial w_{11}} & \cdots & \frac{\partial \delta_{KJ}}{\partial w_{KJ}} \end{pmatrix}$.

According Gale and Nikaido (1965)'s result we know that $\delta(w)$ is globally univalent on $\mathbb{T}_\epsilon$ if $\mathbb{J}_\delta(w)$ is a P-matrix for all $w \in \mathbb{T}_\epsilon$. In the rest of the proof we will show that $\mathbb{J}_\delta(w)$ is indeed a P-matrix whenever Assumption 3 holds.

In the following we will make use extensive use of the following lemma:

Lemma 1. *Under Assumption 1, the following shape restrictions hold:*

$$\frac{\partial s_{kj}}{\partial w_{kl}} \begin{cases} \geq 0, & \text{if } l = j \\ \leq 0, & \text{if } l \in \mathcal{J}_0 \setminus \{j\} \end{cases}$$

Proof.

$$\begin{aligned} s_{kj} &= \mathbb{P}\big(v_{kj} + \epsilon_{ij} \geq v_{kj'} + \epsilon_{ij'} \text{ for all } j' \in \mathcal{J} \cup \{0\} \equiv \mathcal{J}_0\big) \\ &= \mathbb{P}\left(\underbrace{\epsilon_{i0} - \epsilon_{ij}}_{\varepsilon_{ij0}} \leq v_{kj} - v_{k0}, ..., \underbrace{\epsilon_{iJ} - \epsilon_{ij}}_{\varepsilon_{ijJ}} \leq v_{kj} - v_{kJ}\right) \\ &= F_{\varepsilon_{ij0},...,\varepsilon_{ijJ}}(v_{kj} - v_{k0}, ..., v_{kj} - v_{kJ}). \end{aligned}$$

Let $F^{(l)}_{X1,...,X_J}(x_1, ..., x_J) \equiv \frac{\partial}{\partial x_l} F_{X1,...,X_J}(x_1, ..., x_J)$. We have then:

$$\begin{aligned} \frac{\partial s_{kj}}{\partial v_{kl}} &= -F^{(l)}_{\varepsilon_{ij0},...,\varepsilon_{ijJ}}(v_{kj} - v_{k0}, ..., v_{kj} - v_{kJ}) \leq 0, \text{ for } l \neq j, \\ \frac{\partial s_{kj}}{\partial v_{kj}} &= \sum_{l \neq j} F^{(l)}_{\varepsilon_{ij0},...,\varepsilon_{ijJ}}(v_{kj} - v_{k0}, ..., v_{kj} - v_{kJ}) \geq 0, \end{aligned}$$

where both inequalities hold because $F_{\varepsilon_{ij0},...,\varepsilon_{ijJ}}(.)$ is a joint CDF.

□

Definition 2. *Let A be a real square matrix. (i) A is a P-matrix if every principal minor of A is positive, i.e. > 0. (ii) A is said to be a* **positive diagonally dominant** *matrix if there exists a strictly positive vector $d = (d_1, ..., d_n)$ where each $d_i > 0$ such that $d_i A_{ii} > \sum_{j \neq i} d_j |A_{ij}|$.*

According Proposition 1(ii) of Parthasarathy (1983, p.10) any real square matrix that is positive diagonally dominant is a P-matrix. Recall that under Assumption 2, $\mathcal{C}^j \neq \{\emptyset\}$, in fact in our modelling approach $\lambda_j F^j_k(\ell_{\cdot j}(w)) \frac{\mathcal{E}_{kj}(w)}{1+\mathcal{E}_{kj}(w)} = 0 \iff F^j_k(\ell_{\cdot j}(w)) = 0$ for all $w \in \mathbb{T}_\epsilon$, but according Assumption 2, for each $j \in \mathcal{J}$ there exists at least some k such that $F^j_k(\ell_{\cdot j}(w)) > 0$ then $\lambda_j F^j_k(\ell_{\cdot j}(w)) \frac{\mathcal{E}_{kj}(w)}{1+\mathcal{E}_{kj}(w)} > 0$. Under Assumptions 1, and 2, for all $k \in \mathcal{C}^j$ and $j \in \mathcal{J}$, we have

$$\frac{\partial \delta_{kj}}{\partial w_{ml}} = \begin{cases} 1 - \lambda_j \frac{\partial \ell_{kj}(w_{k\cdot})}{\partial w_{kj}} F^j_{kk}(\ell_{\cdot j}(w)) \frac{\mathcal{E}_{kj}(w_{k\cdot})}{1+\mathcal{E}_{kj}(w_{k\cdot})} - \lambda_j F^j_k(\ell_{\cdot j}(w)) \frac{1}{(1+\mathcal{E}_{kj}(w_{k\cdot}))^2} \frac{\partial \mathcal{E}_{kj}(w_{k\cdot})}{\partial w_{kj}}, & \text{if } m = k, l = j \\ -\lambda_j \frac{\partial \ell_{kj}(w_{k\cdot})}{\partial w_{kl}} F^j_{kk}(\ell_{\cdot j}(w)) \frac{\mathcal{E}_{kj}(w_{k\cdot})}{1+\mathcal{E}_{kj}(w_{k\cdot})} - \lambda_j F^j_k(\ell_{\cdot j}(w)) \frac{1}{(1+\mathcal{E}_{kj}(w_{k\cdot}))^2} \frac{\partial \mathcal{E}_{kj}(w_{k\cdot})}{\partial w_{kl}}, & \text{if } m = k, l \neq j \\ -\lambda_j \frac{\partial \ell_{mj}(w_{m\cdot})}{\partial w_{ml}} F^j_{km}(\ell_{\cdot j}(w)) \frac{\mathcal{E}_{kj}(w_{k\cdot})}{1+\mathcal{E}_{kj}(w_{k\cdot})}, & \text{if } m \neq k. \end{cases}$$

for all $(m, l) \in \mathcal{K} \times \mathcal{J}$. Notice that for all $k \in \overline{\mathcal{C}^j} \equiv \mathcal{K} \setminus \mathcal{C}^j$, $j \in \mathcal{J}$, because $F^j_k(\ell_{\cdot j}(w)) = 0$ we have $\frac{\partial \delta_{kj}}{\partial w_{kj}} = 1$ and $\frac{\partial \delta_{kj}}{\partial w_{ml}} = 0$ for $m \neq k$ or $l \neq j$. For all $k \in \mathcal{C}^j$ denote $d_{kj} \equiv w_{kj}/\beta_{kj} > 0$ and for all $k \in \overline{\mathcal{C}^j}$ $d_{kj} = 1$ and this for all $j \in \mathcal{J}$. Let consider two cases:

Case 1: Assumption 3 holds: Under Assumption 3 we have the following sign restriction on $\frac{\partial \delta_{kj}}{\partial w_{ml}}$:

$$\frac{\partial \delta_{kj}}{\partial w_{ml}} = \begin{cases} 1 - \lambda_j \underbrace{\frac{\partial \ell_{kj}(w_{k\cdot})}{\partial w_{kj}}}_{\geq 0} \underbrace{F^j_{kk}(\ell_{\cdot j}(w))}_{\leq 0} \frac{\mathcal{E}_{kj}(w_{k\cdot})}{1+\mathcal{E}_{kj}(w_{k\cdot})} - \lambda_j F^j_k(\ell_{\cdot j}(w)) \frac{1}{(1+\mathcal{E}_{kj}(w_{k\cdot}))^2} \underbrace{\frac{\partial \mathcal{E}_{kj}(w_{k\cdot})}{\partial w_{kj}}}_{\leq 0} > 0, \text{ if } m = k, l = j \\ -\lambda_j \underbrace{\frac{\partial \ell_{kj}(w_{k\cdot})}{\partial w_{kl}}}_{\leq 0} \underbrace{F^j_{kk}(\ell_{\cdot j}(w))}_{\leq 0} \frac{\mathcal{E}_{kj}(w_{k\cdot})}{1+\mathcal{E}_{kj}(w_{k\cdot})} - \lambda_j F^j_k(\ell_{\cdot j}(w)) \frac{1}{(1+\mathcal{E}_{kj}(w_{k\cdot}))^2} \underbrace{\frac{\partial \mathcal{E}_{kj}(w_{k\cdot})}{\partial w_{kl}}}_{\geq 0} \leq 0, \text{ if } m = k, l \neq j, \\ -\lambda_j \frac{\partial \ell_{mj}(w_{m\cdot})}{\partial w_{ml}} \underbrace{F^j_{km}(\ell_{\cdot j}(w))}_{=0} \frac{\mathcal{E}_{kj}(w_{k\cdot})}{1+\mathcal{E}_{kj}(w_{k\cdot})} = 0, \text{ if } m \neq k. \end{cases}$$

Therefore, for all $k \in \mathcal{C}^j$ and $j \in \mathcal{J}$, we can show that

$$\begin{aligned} &\frac{w_{kj}}{\beta_{kj}} \frac{\partial \delta_{kj}}{\partial w_{kj}} - \sum_{m \neq k \text{ or } l \neq j} \frac{w_{ml}}{\beta_{ml}} \left| \frac{\partial \delta_{kj}}{\partial w_{ml}} \right| = \\ &\qquad \underbrace{\frac{w_{kj}}{\beta_{kj}}}_{>0} - \lambda_j \underbrace{\left[\frac{w_{kj}}{\beta_{kj}} \frac{\partial \ell_{kj}(w_{k\cdot})}{\partial w_{kj}} + \sum_{l \neq j} \frac{w_{kl}}{\beta_{kl}} \frac{\partial \ell_{kj}(w_{k\cdot})}{\partial w_{kl}} \right]}_{m_k \sum_{l \in \mathcal{J}} \frac{\partial s_{kj}(w_{k\cdot})}{\partial v_{kl}} = -m_k \frac{\partial s_{kj}(w_{k\cdot})}{\partial v_{k0}} \geq 0} \underbrace{F^j_{kk}(\ell_{\cdot j}(w))}_{\leq 0} \frac{\mathcal{E}_{kj}(w_{k\cdot})}{1+\mathcal{E}_{kj}(w_{k\cdot})} \\ &\qquad - \lambda_j \underbrace{F^j_k(\ell_{\cdot j}(w))}_{\geq 0} \frac{1}{(1+\mathcal{E}_{kj}(w_{k\cdot}))^2} \underbrace{\left[\frac{w_{kj}}{\beta_{kj}} \frac{\partial \mathcal{E}_{kj}}{\partial w_{kj}} + \sum_{l \neq j} \frac{w_{kl}}{\beta_{kl}} \frac{\partial \mathcal{E}_{kj}(w_{k\cdot})}{\partial w_{kl}} \right]}_{\sum_{l \in \mathcal{J}} \frac{\partial \mathcal{E}_{kj}(w_{k\cdot})}{\partial v_{kl}} = -\frac{\partial \mathcal{E}_{kj}}{\partial v_{k0}} \leq 0} > 0. \end{aligned} \tag{B.3}$$

All the sign restrictions hold under Assumption 3 holds. Two main non-obvious points in the previous inequality are the following equalities $\sum_{l \in \mathcal{J}_0} \frac{\partial s_{kj}(w_{k\cdot})}{\partial v_{kl}} = 0$ and $\sum_{l \in \mathcal{J}_0} \frac{\partial \mathcal{E}_{kj}(w_{k\cdot})}{\partial v_{kl}} = 0$. The trick behind these equalities is the fact that an increase of all mean gross utility $v_{k\cdot}$ does not affect the share s_{kj} as remarked by Berry (1994, page 267). The same argument applies also to the elasticity which justifies the second equality. Moreover for all $k \in \overline{\mathcal{C}^j}$, and $j \in \mathcal{J}$, $d_{kj} \frac{\partial \delta_{kj}}{\partial w_{kj}} - \sum_{m \neq k \text{ or } l \neq j} d_{ml} \left| \frac{\partial \delta_{kj}}{\partial w_{ml}} \right| > 0$ trivially holds. Therefore, $\mathbb{J}_\delta(w)$ is indeed a P-matrix for all $w \in \mathbb{T}_\epsilon$, and then $\delta(w)$ is globally univalent on $\mathbb{T}_\epsilon$, which complete the proof.

Case 2: Assumption 3 (i) holds: In such a context we can show that

$$\begin{aligned} &\frac{w_{kj}}{\beta_{kj}} \frac{\partial \delta_{kj}}{\partial w_{kj}} - \sum_{m \neq k \text{ or } l \neq j} \frac{w_{ml}}{\beta_{ml}} \left| \frac{\partial \delta_{kj}}{\partial w_{ml}} \right| = \\ &\qquad \frac{w_{kj}}{\beta_{kj}} + \lambda_j \sum_{m \neq k} \underbrace{\left[-\frac{w_{mj}}{\beta_{mj}} \frac{\partial \ell_{mj}(w_{m\cdot})}{\partial w_{mj}} + \sum_{l \neq j} \frac{w_{ml}}{\beta_{ml}} \frac{\partial \ell_{mj}(w_{m\cdot})}{\partial w_{ml}} \right]}_{-\frac{\partial \ell_{mj}}{\partial v_{m0}} - 2\frac{\partial \ell_{mj}}{\partial v_{mj}}} \left| F^j_{km}(\ell_{\cdot j}(w)) \right| \frac{\mathcal{E}_{kj}(w_{k\cdot})}{1+\mathcal{E}_{kj}(w_{k\cdot})} \\ &\qquad - \lambda_j \underbrace{\left[\frac{w_{kj}}{\beta_{kj}} \frac{\partial \ell_{kj}(w_{k\cdot})}{\partial w_{kj}} + \sum_{l \neq j} \frac{w_{kl}}{\beta_{kl}} \frac{\partial \ell_{kj}(w_{k\cdot})}{\partial w_{kl}} \right]}_{m_k \sum_{l \in \mathcal{J}} \frac{\partial s_{kj}(w_{k\cdot})}{\partial v_{kl}} = -m_k \frac{\partial s_{kj}(w_{k\cdot})}{\partial v_{k0}} \geq 0} \underbrace{F^j_{kk}(\ell_{\cdot j}(w))}_{\leq 0} \frac{\mathcal{E}_{kj}(w_{k\cdot})}{1+\mathcal{E}_{kj}(w_{k\cdot})} \\ &\qquad - \lambda_j \underbrace{F^j_k(\ell_{\cdot j}(w))}_{\geq 0} \frac{1}{(1+\mathcal{E}_{kj}(w_{k\cdot}))^2} \underbrace{\left[\frac{w_{kj}}{\beta_{kj}} \frac{\partial \mathcal{E}_{kj}}{\partial w_{kj}} + \sum_{l \neq j} \frac{w_{kl}}{\beta_{kl}} \frac{\partial \mathcal{E}_{kj}(w_{k\cdot})}{\partial w_{kl}} \right]}_{\sum_{l \in \mathcal{J}} \frac{\partial \mathcal{E}_{kj}(w_{k\cdot})}{\partial v_{kl}} = -\frac{\partial \mathcal{E}_{kj}}{\partial v_{k0}} \leq 0}. \end{aligned}$$

Notice that the second term after the equality holds because, as discussed earlier, we have $\sum_{l\in\mathcal{J}}\frac{\partial s_{mj}(w_{m\cdot})}{\partial v_{ml}} = -\frac{\partial s_{mj}(w_{m\cdot})}{\partial v_{m0}}$. Therefore, we can write:

$$\frac{w_{kj}}{\beta_{kj}}\frac{\partial\delta_{kj}}{\partial w_{kj}} - \sum_{m\neq k \text{ or } l\neq j}\frac{w_{ml}}{\beta_{ml}}\left|\frac{\partial\delta_{kj}}{\partial w_{ml}}\right| = \frac{w_{kj}}{\beta_{kj}} + \lambda_j\Bigg\{-\sum_{m\neq k}\left[\frac{\partial\ell_{mj}(w_{m\cdot})}{\partial v_{m0}} + 2\frac{\partial\ell_{mj}(w_{m\cdot})}{\partial v_{mj}}\right]\left|F^j_{km}(\ell_{\cdot j}(w))\right|$$
$$+\frac{\partial\ell_{kj}(w_{k\cdot})}{\partial v_{k0}}F^j_{kk}(\ell_{\cdot j}(w)) + F^j_k(\ell_{\cdot j}(w))\frac{1}{(1+\mathcal{E}_{kj}(w_{k\cdot}))\mathcal{E}_{kj}(w_{k\cdot})}\frac{\partial\mathcal{E}_{kj}}{\partial v_{k0}}\Bigg\}\times\frac{\mathcal{E}_{kj}(w_{k\cdot})}{1+\mathcal{E}_{kj}(w_{k\cdot})}.$$

As can be seen, without additive separability in the production function the equilibrium can be unique if the RHS of the latter equality is positive. A sufficient condition for it is that

$$\Bigg\{-\sum_{m\neq k}\left[\frac{\partial\ell_{mj}(w_{m\cdot})}{\partial v_{m0}} + 2\frac{\partial\ell_{mj}(w_{m\cdot})}{\partial v_{mj}}\right]\left|F^j_{km}(\ell_{\cdot j}(w))\right| + \frac{\partial\ell_{kj}(w_{k\cdot})}{\partial v_{k0}}F^j_{kk}(\ell_{\cdot j}(w)) +$$
$$+F^j_k(\ell_{\cdot j}(w))\frac{1}{(1+\mathcal{E}_{kj}(w_{k\cdot}))\mathcal{E}_{kj}(w_{k\cdot})}\frac{\partial\mathcal{E}_{kj}}{\partial v_{k0}}\Bigg\}\geq 0 \tag{B.4}$$

for all $w\in\mathbb{T}_\epsilon$.

B.3. **Proof of Proposition 1.**

Lemma 2. *Under Assumptions 1, 2, and 3, $\delta(w)$ is generalized nonlinear diagonally dominant on $\mathbb{T}_\epsilon$.*

Proof. All partial derivative of $\delta(w)$ exists and are continuous. Let's $\mathbb{J}_\delta(w)\equiv\delta(w)'$ be its Jacobian matrix which is continuous on $\mathbb{T}_\epsilon$. $\delta(w)$ is Frèchet-differentiable on $\mathbb{T}_\epsilon$ then it is Gâteaux-differentiable on $\mathbb{T}_\epsilon$ which is a convex compact subset of $\mathbb{R}_{KJ}$. In the case 1 of the Proof of Theorem 2, we show that $\mathbb{J}_\delta(w)$ is *a generalized diagonally dominant matrix* in the language of Gan et al. (2006) and this for all $w\in\mathbb{T}_\epsilon$. The proof is complete once we invoke Theorem 8 of Gan et al. (2006).

□

Lemma 3. *Under Assumptions 1, 2, and 3,*

For any $w\in\mathbb{T}_\epsilon$, and $kj=1,...,KJ$ the following equation in x_{kj}:
$\psi(x_{kj},w_{-kj})\equiv\delta_{kj}(w_{11},...,w_{1J},...,w_{k,j-1},x_{kj},w_{k,j+1},...,w_{KJ})=0$ as a unique solution x^*_{kj}.

Proof. In the case 1 of the Proof of Theorem 2, we show that $\frac{\partial\psi(x_{kj},w_{-kj})}{\partial x_{ij}}\geq 1>0$, then $\psi(x_{kj},w_{-kj})$ is strictly increasing in x_{kj} for any $w_{-kj}\in\mathbb{T}_\epsilon$. In addition, as can be seen in the proof of Theorem 1, $\psi(\underline{\epsilon}_{kj},w_{-kj})\leq 0\leq\psi(\bar{\lambda}\bar{F}'-\bar{\epsilon}_{kj},w_{-kj})$ for for any $w_{-kj}\in\mathbb{T}_\epsilon$. This completes the proof. □

Under Assumptions 1, 2, and 3 Lemmata 2, and 3 hold, then we could invoke Theorem 18 of Frommer (1991). Remark that both underrelaxed Gauss-siedel and Jacobi iteration are special cases of the asynchronous iterative methods discussed in Frommer (1991) Theorem 18. This complete the Proof of Proposition 1.

B.4. **Proof of Proposition 2.**

Under Assumptions 1, 2, and 3, we proved that we have an unique equilibrium w^{eq} such that $w^{eq}=B(w^{eq})$. For sake of simplicity let us ignore the upper-script eq in the rest of the proof. By the Implicit Function Theorem we have:

$$\frac{dw}{dw_{k0}} = \mathbb{J}_\delta^{-1}(w)\frac{\partial B(w)}{\partial w_{k0}},$$
$$\frac{dw}{d\gamma_{kl}} = \mathbb{J}_\delta^{-1}(w)\frac{\partial B(w)}{\partial \gamma_{kl}},$$
$$\frac{dw}{d\theta_l} = \mathbb{J}_\delta^{-1}(w)\frac{\partial B(w)}{\partial \theta_l}.$$

Under Assumption 3, $\mathbb{J}_\delta(w)$ is a block diagonal matrix, more precisely it can be written $\underset{(KJ\times KJ)}{\mathbb{J}_\delta(w)} = \begin{pmatrix} \mathbb{J}_{\delta,1\cdot}(w) & 0 & \cdots & 0 \\ 0 & \mathbb{J}_{\delta,2\cdot}(w) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbb{J}_{\delta,K\cdot}(w) \end{pmatrix}$, where $\underset{(J\times J)}{\mathbb{J}_{\delta,k\cdot}(w)} = \begin{pmatrix} \frac{\partial \delta_{k1}}{\partial w_{k1}} & \cdots & \frac{\partial \delta_{k1}}{\partial w_{kJ}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \delta_{kJ}}{\partial w_{k1}} & \cdots & \frac{\partial \delta_{kJ}}{\partial w_{kJ}} \end{pmatrix}$. The case 1 of the Proof of Theorem 2, shows that each $\mathbb{J}_{\delta,k\cdot}(w)$ for $k \in \mathcal{K}$ is positive diagonally dominant, therefore its inverse exists and then we have, $\underset{(KJ\times KJ)}{\mathbb{J}_\delta^{-1}(w)} = \begin{pmatrix} \mathbb{J}_{\delta,1\cdot}^{-1}(w) & 0 & \cdots & 0 \\ 0 & \mathbb{J}_{\delta,2\cdot}^{-1}(w) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbb{J}_{\delta,K\cdot}^{-1}(w) \end{pmatrix}$. We then have $\frac{dw_{m\cdot}}{dw_{k0}} = \mathbb{J}_{\delta,m\cdot}^{-1}(w)\frac{\partial B_{m\cdot}(w)}{\partial w_{k0}}$ where $w_{m\cdot} = \begin{pmatrix} w_{m1} \\ \vdots \\ w_{mJ} \end{pmatrix}$, and $B_{m\cdot}(w) = \begin{pmatrix} B_{m1}(w) \\ \vdots \\ B_{mJ}(w) \end{pmatrix}$. Our derived bounds come from the linear algebra results on M-matrices and inverse M-matrices, i.e. Carlson and Markham (1979); Fiedler and Pták (1962). In fact, case 1 of the Proof of Theorem 2, shows that any $\mathbb{J}_{\delta,k\cdot}(w)$ for $k \in \mathcal{K}$ is positive diagonally dominant and have non-positive off diagonal elements. Then, $\mathbb{J}_{\delta,k\cdot}(w)$, and $\mathbb{J}_\delta(w)$ are M Matrices. Our proofs widely use the result (4.2) of Fiedler and Pták (1962), which states that if A and B are two M matrices such that $A \leqq B$, then $A^{-1} \geqq B^{-1} \geqq 0$. Let's denote by DA the diagonal matrix formed by the diagonal elements of the matrix A. Under Assumption 3, we have $\mathbb{J}_{\delta,k\cdot}(w) \le D\mathbb{J}_{\delta,k\cdot}(w) \Rightarrow \mathbb{J}_{\delta,k\cdot}^{-1}(w) \ge [D\mathbb{J}_{\delta,k\cdot}(w)]^{-1} \Rightarrow \mathbb{J}_{\delta,k\cdot}^{-1}(w)\frac{\partial B_{k\cdot}(w)}{\partial w_{k0}} \ge [D\mathbb{J}_{\delta,k\cdot}(w)]^{-1}\frac{\partial B_{k\cdot}(w)}{\partial w_{k0}}$ where the last inequality holds since $\frac{\partial B_{kj}(w)}{\partial w_{k0}} \ge 0$ under Assumption 3.

It follows from the latter inequality that:

$$\frac{\partial w_{kj}}{\partial w_{k0}} \ge \frac{w_{kj}}{w_{k0}}\frac{\psi_{k,j0}}{1-\psi_{k,jj}} \ge 0$$

where $\psi_{k,jl} = \left(\frac{w_{kl}}{\ell_{kj}}\frac{\partial \ell_{kj}(w_{k\cdot})}{\partial w_{kl}}\left(\frac{F^j_{kk}}{F^j_k}\ell_{kj}\right) + \frac{1}{(1+\mathcal{E}_{kj}(w_{k\cdot}))}\frac{w_{kl}}{\mathcal{E}_{kj}(w_{k\cdot})}\frac{\partial \mathcal{E}_{kj}(w_{k\cdot})}{\partial w_{kl}}\right)$. This latter inequality becomes evident as soon as you remark that: $\frac{\partial \delta_{kj}}{\partial w_{kl}} \begin{cases} -\left(\frac{w_{kj}}{w_{kl}}\right)\psi_{k,jl} \text{ if } j \neq l \\ 1-\psi_{k,jl} \text{ if } j = l \end{cases}$. This proves the first set of bounds.

Second, for $a_{ll} > 0$ and $a_{jl} \leq 0$ when $j \neq l$ it can be shown that

$$H^{-1}(a_{\cdot\cdot}) \equiv \begin{pmatrix} a_{11} & 0 & \cdots & 0 & a_{1l} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & 0 & a_{ll} & 0 & \cdots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & a_{l+1,l+1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & a_{J,J} \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} 1/a_{11} & 0 & \cdots & 0 & -a_{1l}/a_{11}a_{ll} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & 0 & 1/a_{ll} & 0 & \cdots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & 1/a_{l+1,l+1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1/a_{J,J} \end{pmatrix},$$

$\frac{\partial B_{k\cdot}(w)}{\partial \theta_l} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ B_{kl}(w)/\theta_l \\ 0 \\ \vdots \\ 0 \end{pmatrix} \geq 0$. For $a_{jl} \equiv \frac{\partial \delta_{kj}}{\partial w_{kl}}$ we have $\mathbb{J}_{\delta,k\cdot}(w) \leq H\left(\frac{\partial \delta_{k\cdot}}{\partial w_{k\cdot}}\right) \Rightarrow \mathbb{J}^{-1}_{\delta,k\cdot}(w) \geq \left[H\left(\frac{\partial \delta_{k\cdot}}{\partial w_{k\cdot}}\right)\right]^{-1} \Rightarrow$ $\mathbb{J}^{-1}_{\delta,k\cdot}(w)\frac{\partial B_{k\cdot}(w)}{\partial \theta_l} \geq \left[H\left(\frac{\partial \delta_{k\cdot}}{\partial w_{k\cdot}}\right)\right]^{-1}\frac{\partial B_{k\cdot}(w)}{\partial \theta_l}$. The latter inequality implies that fo $j \leq l$ we have:

$$\frac{\partial w_{kj}}{\partial \theta_l} \begin{cases} \geq -\dfrac{\frac{\partial \delta_{kj}}{\partial w_{kl}}}{\frac{\partial \delta_{kj}}{\partial w_{kj}}\frac{\partial \delta_{kl}}{\partial w_{kl}}}\dfrac{B_{kl}(w)}{\theta_l} = \dfrac{w_{kj}\psi_{k,jl}}{\theta_l(1-\psi_{k,jj})(1-\psi_{k,ll})} \geq 0 \text{ if } j < l \\ \geq \dfrac{1}{\frac{\partial \delta_{kl}}{\partial w_{kl}}}\dfrac{B_{kl}(w)}{\theta_l} = \dfrac{w_{kl}}{\theta_l(1-\psi_{k,ll})} > 0, \text{ if } j = l. \text{ otherwise.} \end{cases} \tag{B.5}$$

For $j < l$, we can follow the same process by considering H as a lower triangular matrix. The exact same proof holds for $\frac{\partial w_{kj}}{\partial \theta_l}$. This completes the proof.

Special case: Duopsony. In this special case, we could have a passthrough formula that will hold at equality. This will allow us to have an intuition of the shock transmission from a firm j to a firm l. Recall that $\frac{dw_{m\cdot}}{dw_{k0}} = \mathbb{J}^{-1}_{\delta,m\cdot}(w)\frac{\partial B_{m\cdot}(w)}{\partial w_{k0}}$, and $\frac{\partial \delta_{kj}}{\partial w_{kl}} = -\left(\frac{w_{kj}}{w_{kl}}\right)\psi_{k,jl}$ for $l \neq j$.

Now, consider that $\mathcal{J} = \{j, l\}$. In this special case the inverse of the Jacobian matrix is given by: $(\mathbb{J}_{\delta,k\cdot}(w))^{-1} = \begin{pmatrix} \frac{\partial \delta_{kj}}{\partial w_{kj}} & \frac{\partial \delta_{kj}}{\partial w_{kl}} \\ \frac{\partial \delta_{kl}}{\partial w_{kj}} & \frac{\partial \delta_{kl}}{\partial w_{kl}} \end{pmatrix}^{-1} = \frac{1}{(1-\psi_{k,jj})(1-\psi_{k,ll})-\psi_{k,jl}\psi_{k,lj}} \begin{pmatrix} (1-\psi_{k,ll}) & \left(\frac{w_{kj}}{w_{kl}}\right)\psi_{k,jl} \\ \left(\frac{w_{kl}}{w_{kj}}\right)\psi_{k,lj} & (1-\psi_{k,jj}) \end{pmatrix}$. Then we can easily derive the following:

$$\frac{w_{k0}}{w_{kj}} \frac{\partial w_{kj}}{\partial w_{k0}} = \frac{(1-\psi_{k,ll})\psi_{k,j0} + \psi_{k,jl}\psi_{k,l0}}{(1-\psi_{k,jj})(1-\psi_{k,ll}) - \psi_{k,jl}\psi_{k,lj}} \geq 0 \tag{B.6}$$

$$\frac{u_{kl}}{w_{kj}} \frac{\partial w_{kj}}{\partial u_{kl}} = \frac{(1-\psi_{k,ll})\phi_{k,jl} + \psi_{k,jl}\phi_{k,ll}}{(1-\psi_{k,jj})(1-\psi_{k,ll}) - \psi_{k,jl}\psi_{k,lj}} \gtreqless 0 \tag{B.7}$$

$$\frac{u_{kl}}{w_{kl}} \frac{\partial w_{kl}}{\partial u_{kl}} = \frac{(1-\psi_{k,jj})\phi_{k,ll} + \psi_{k,lj}\phi_{k,jl}}{(1-\psi_{k,jj})(1-\psi_{k,ll}) - \psi_{k,jl}\psi_{k,lj}} \gtreqless 0 \tag{B.8}$$

$$\frac{\theta_l}{w_{kj}} \frac{\partial w_{kj}}{\partial \theta_l} = \frac{\psi_{k,jl}}{(1-\psi_{k,jj})(1-\psi_{k,ll}) - \psi_{k,jl}\psi_{k,lj}} \geq 0 \tag{B.9}$$

$$\frac{\theta_l}{w_{kj}} \frac{\partial w_{kj}}{\partial \theta_l} = \frac{(1-\psi_{k,jj})}{(1-\psi_{k,jj})(1-\psi_{k,ll}) - \psi_{k,jl}\psi_{k,lj}} \geq 0 \tag{B.10}$$

where the signs restrictions hold, because $\psi_{k,jl}, \phi_{k,jl} \geq 0$ for $l \neq j$, and $\psi_{k,ll}, \phi_{k,ll} \leq 0$ with $\phi_{k,jl} = \left(\frac{u_{kl}}{\ell_{kj}} \frac{\partial \ell_{kj}(w_{k\cdot})}{\partial u_{kl}} \left(\frac{F^j_{kk}}{F^j_k} \ell_{kj} \right) + \frac{1}{(1+\mathcal{E}_{kj}(w_{k\cdot}))} \frac{u_{kl}}{\mathcal{E}_{kj}(w_{k\cdot})} \frac{\partial \mathcal{E}_{kj}(w_{k\cdot})}{\partial u_{kl}} \right)$

Appendix C. Data and Sample Description

Our data consists of several administrative registers provided by Statistics Denmark for the years 2001-2019. These include annual cross-section data from the Danish register-based, matched employer-employee dataset IDA (Integrated Database for Labor Market Research) and other annual datasets, divided into IDAN, IDAS, and IDAP. The datasets are linked by individual identifiers for persons, firms, and establishments. Table C.1 lists the relevant datasets and details.

Table C.1. Data Description (Datasets and Variables).

Category	**Register**	**Variables**
workers	IDAN (jobs yearly panel)	firm and establishment indicator, establishment location, yearly earnings, hours worked, share of the year worked, type of job (primary, secondary), type of job (part-time/full-time), type of job (occupation, DISCO code)
not employed	BEF (population register) IDAN	We classify as not employed all individuals in the relevant age groups who are not recorded in IDAN.
unemployed	IND (income dataset, individual yearly panel), IDAP (worker dataset, individual yearly panel)	unemployment benefits, duration of unemployment
firms and establishments	FIRM, IDAS (workplace panel)	firm revenue, sector of industry (5-digit industry classification based on NACE rev. 2), establishment location (municipality)
k-groups	UDDA (education panel), BEF (individual yearly panel)	age, highest acquired education, sex
commuting zones	Eckert et al. (2022) (available on Fabian Eckert website)	commuting zone (link to municipality)

We restrict the dataset to individuals between 26 and 60 years of age who work full-time as employees in the private sector whose job is linked to a physical establishment. We drop individuals employed in the financial sector; firms in the financial sector are not required to report revenue data and very few do. Details on data and sample selection are in table C.2. In total, our dataset consists of $12,742,746$ individual-year combinations. Our sample construction selects the data in a few important ways: The full population of salaried jobs in Denmark in 2001-2019 is 49.3 percent female. This goes down to 35.8 percent when we drop the public sector and further to 31.8 percent when we exclude the financial sector and non-full-time jobs. Workers in the private-sector with full-time jobs are on average one year older than the full worker population, and have average yearly earnings of $71,491$ USD, higher than the full-worker-population average of $42,867$ USD.

Table C.2. Worker Sample Selection.

	step	observations	share in public sector	share in financial sector	share full-time	share female	avg. age	avg. yearly earnings (2022 USD)
1	All salaried jobs in Denmark between 2001 and 2019	76,869,608						
2	Keep jobs held by workers in relevant k-groups	50,263,511	0.229	0.024	0.437	0.493	42.454	42,867
3	Keep jobs with market information (primary jobs)	32,486,151	0.355	0.037	0.648	0.487	42.964	56,389
4	Drop workers in small commuting zones	32,106,644	0.354	0.037	0.768	0.487	42.943	56,474
5	Drop jobs with no earnings or hours	32,094,227	0.354	0.037	0.648	0.487	42.944	56,493
6	Drop public sector jobs	20,719,775		0.057	0.660	0.358	42.482	59,641
7	Drop financial sector jobs	19,538,794			0.653	0.349	42.425	58,296
8	Keep full-time, highest-paying jobs	**12,742,746**				0.318	43.518	71,491
10	Only period 2004-2016	8,614,260						

Find a detailed description of the selection steps below:

(1) This step excludes self employed and employers, as well as their spouses if their main source of income is from assisting the spouse's enterprise; it includes all other types of jobs.
(2) This step drops workers not appearing in the population registers, younger and older workers, as well as workers with no education status recorded (this applies mostly to immigrant workers). Therefore, this step excludes jobs held by workers not resident in Denmark.
(3) This step drops jobs without real establishment code, i.e., all non-primary jobs and primary jobs with missing or fictitious establishment code. Primary jobs are the most important connection to the labor market (longest employment period and largest ATP payments). Workers with fictitious workplaces (establishment nr. = 0) are those who cannot be linked to any of the employer's registered workplaces, either because they work from home or in various workplaces (such as cleaners, home nurses). Workers with no workplace (establishment nr. = .) are those with multiple workplaces for which one unique workplace cannot be identified. In 2,491,168 instances, where the establishment information is missing only in one year during a continuous employment spell at the same firm, we impute it.
(4) Drop jobs in establishments in the islands of Christiansœ, Bornholm, Samsœ, and Æro.
(5) Drop jobs with no information on earnings or hours
(6) Drop if the sector of industry of the employer is one of the following nacee-2 codes {O,P,Q,T,U,X}.
(7) Drop if the sector of industry of the employer is nacee-2 code K (this sector has an extreme underreporting of revenue data).
(8) We define full-time jobs as jobs with weekly schedule of 30 hours or more.

We denote establishments with the subscript j, time (years) with the subscript t, and worker type (or k-groups) with the subscript k. Worker types are divided by sex (male or female) age group (26-35, 36-50, 51-60) and education level (completed or not tertiary education). We collapse the individual-level dataset at the (k, j, t) level leading to $4,487,628$ observations. We restrict the estimation dataset to only establishments that have no missing values for any of the key variables. Table C.3 details the sample selection process.

The key variables we use in the estimation are:

- w_{kjt}: mean earnings by k-group, establishment, year
- w_{k0t}: mean non-employment income by k-group, year

Table C.3. Establishment Sample Selection and Construction of the Estimation Dataset.

	step	total observations	unique establishments
1	collapse at the kgroup-establishment-year (k, j, t) level	4,487,628	259,195
2	merge revenue data (firm, year)	-	-
3	add share of non-employed/unemployed and average unemployment income	-	-
4	drop observations with wage bill to revenue ratio above 80% (drops all observations with missing revenue)	4,054,235	238,299
	keep observations for firms that appear at least once in the estimation dataset	**3,069,490**	**63,525**
5	create estimation variables	-	-
6	keep observations in 2004-2017 to accommodate for long run lags $(x_{jkt+2} - x_{jkt-3})$ and data break	2,268,523	-
7	drop firms/k-groups with not enough longevity to allow for computing short-run lags $(x_{jkt} - x_{jkt-1})$	2,318,335	-
8	drop firms/k-groups with not enough longevity to allow for computing long-run lags $(x_{jkt+2} - x_{jkt-3})$	1,914,366	-
9	drop firms employing only one k-group (necessary for the second instrument)	**1,101,541**	**63,525**

Start with panel of selected workers in years 2001-2019. Variables: full-time-equivalent, earnings, k-group (sex, age, education), local market (commuting zone, industry), firm, establishment, year (12,742,746 individuals).

- s_{kjt} and $s_{kj|gt}$: employment shares, by k-group, establishment, year, overall and by market g (inside shares)
- $s_{\sim kj|gt}$: sum of the inside shares for all other labor types employed by establishment j, by k-group, year, market
- R_{jt}: establishment-level revenue by year, obtained allocating firm revenue across establishments in proportion to their wage bills

Appendix D. Appendix Figures and Tables

Table D.1. Establishment characteristics, by commuting zone (full sample)

	n. unique estab.	n. estab. per firm		n. of workers per estab.		n. of k-groups per estab.		estab. revenue (1,000 UDS)		average wage (USD)	
commuting zone		mean	st. dev.	mean	st. dev.	mean	st. dev.	mean	st. dev.	mean	st. dev.
1. North and East Zealand (Copenhagen)	92,763	1.225	3.535	8.454	40.626	2.685	2.356	6,171	61,366	65,148	36,819
2. West and South Zealand (Slagelse)	10,714	1.229	3.378	5.816	33.274	2.348	1.897	3,941	55,333	55,326	15,962
3. West and South Zealand (Køge)	11,953	1.205	3.255	5.715	19.334	2.383	1.929	3,533	22,430	55,888	17,122
4. West and South Zealand (Nykøbing Falster)	4,432	1.249	3.408	5.294	14.390	2.316	1.794	2,729	10,837	51,730	13,878
7. Fyn (Odense)	18,870	1.251	3.679	7.285	24.103	2.686	2.251	4,829	29,984	56,571	26,792
8. Fyn (Svendborg)	2,927	1.183	2.346	4.953	9.919	2.400	1.917	2,820	8,953	54,654	17,221
9. South Jutland (Sønderborg)	5,721	1.224	2.299	8.191	48.528	2.613	2.162	5,614	31,400	54,921	16,691
10. South Jutland (Ribe)	2,041	1.137	1.874	5.554	17.850	2.298	1.879	4,179	22,629	52,261	13,967
11. South Jutland (Kolding)	9,586	1.285	4.333	7.323	19.109	2.727	2.280	4,924	17,372	56,779	17,734
12. Mid-South Jutland (Vejle)	14,569	1.223	3.707	7.820	45.272	2.680	2.258	6,017	58,046	57,835	21,745
13. South-West Jutland (Esbjerg)	10,559	1.218	3.293	6.981	22.509	2.590	2.167	5,419	58,484	55,862	16,837
14. West Jutland (Herning)	9,536	1.233	3.943	7.040	22.462	2.605	2.156	4,583	22,913	55,664	15,332
15. North-West Jutland (Thisted)	2,135	1.172	2.080	6.329	21.196	2.416	1.975	4,009	15,606	54,166	13,972
16. East Jutland (Aarhus)	31,828	1.232	3.362	7.399	24.617	2.678	2.271	5,160	53,258	59,101	22,934
17. Mid-North Jutland (Viborg)	7,988	1.169	2.632	6.901	47.707	2.493	2.077	4,071	26,117	54,906	15,958
19. North Jutland (Aalborg)	23,573	1.232	3.772	6.523	21.000	2.520	2.115	4,499	49,905	55,542	18,252
All of Denmark	259,195	1.227	3.494	7.414	33.071	2.611	2.223	5,198	49,994	59,311	27,048

Source: Administrative registers, Statistics Denmark. Full population of private sector establishments in Denmark (step 1 in table C.3). Commuting zones computed for 2005 by Eckert et al. (2022), largest city in parentheses. We drop six small islands and we merge Aalborg and Frederikshavn. Revenue and average wage at the firm in 2022 USD.

Table D.2. Establishment characteristics, by commuting zone (estimation sample, all years)

	n. unique estab.	n. estab. per firm		n. of workers per estab.		n. of k-groups per estab.		estab. revenue (1,000 UDS)		average wage (USD)	
commuting zone		mean	st. dev.	mean	st. dev.	mean	st. dev.	mean	st. dev.	mean	st. dev.
1. North and East Zealand (Copenhagen)	20,358	1.204	3.052	13.603	54.919	3.672	2.581	11,148	71,722	68,275	24,494
2. West and South Zealand (Slagelse)	2,586	1.257	4.753	9.008	46.099	3.106	2.087	6,915	77,428	57,408	13,925
3. West and South Zealand (Køge)	2,827	1.205	2.702	9.000	26.336	3.212	2.113	6,138	31,361	58,367	15,116
4. West and South Zealand (Nykøbing Falster)	1,099	1.272	3.142	7.981	18.888	3.028	1.954	4,528	14,376	53,726	12,963
7. Fyn (Odense)	4,904	1.220	2.536	11.125	31.530	3.575	2.438	7,969	39,128	58,870	16,999
8. Fyn (Svendborg)	751	1.146	1.402	7.356	12.324	3.189	2.086	4,691	11,938	56,857	14,928
9. South Jutland (Sønderborg)	1,554	1.238	3.337	12.882	65.503	3.433	2.352	9,352	41,824	57,073	14,425
10. South Jutland (Ribe)	512	1.139	1.351	9.010	24.495	3.129	2.118	7,210	31,337	54,684	12,545
11. South Jutland (Kolding)	2,636	1.263	2.919	11.245	24.572	3.613	2.485	8,000	22,656	59,639	16,070
12. Mid-South Jutland (Vejle)	3,934	1.209	2.927	12.184	60.546	3.587	2.456	10,022	78,264	60,382	18,023
13. South-West Jutland (Esbjerg)	2,915	1.207	2.043	10.648	28.233	3.445	2.362	8,952	79,013	58,512	15,138
14. West Jutland (Herning)	2,672	1.199	3.521	10.817	29.165	3.464	2.344	7,453	30,365	57,933	13,439
15. North-West Jutland (Thisted)	585	1.205	3.829	9.958	28.068	3.217	2.174	6,650	20,664	56,651	12,735
16. East Jutland (Aarhus)	8,203	1.248	3.359	11.303	31.179	3.588	2.456	8,625	72,640	61,478	16,908
17. Mid-North Jutland (Viborg)	2,092	1.191	4.713	10.737	65.201	3.349	2.254	6,614	28,985	57,435	15,628
19. North Jutland (Aalborg)	5,897	1.236	3.810	10.202	27.552	3.412	2.330	7,560	67,593	57,936	16,245
All of Denmark	63,525	1.219	3.240	11.591	44.041	3.515	2.427	8,909	62,962	61,787	19,573

Source: Administrative registers, Statistics Denmark. Restricted sample of establishments with no missing values for the key estimation variables (step 5 in table C.3). Commuting zones computed for 2005 by Eckert et al. (2022), largest city in parentheses. We drop six small islands and we merge Aalborg and Frederikshavn. Revenue and average wage at the firm in 2022 USD.

Table D.3. Establishment characteristics, by industry (full sample)

	n. unique estab.	n. estab. per firm		n. of workers per estab.		n. of k-groups per estab.		estab. revenue (1,000 UDS)		average wage (USD)	
commuting zone		mean	st. dev.	mean	st. dev.	mean	st. dev.	mean	st. dev.	mean	st. dev.
A. Agriculture, forestry, and fishery	13,486	1.042	0.711	2.302	4.045	1.643	1.246	1,720	2,909	48,810	13,767
B. Mining and quarrying	425	1.690	3.088	13.872	62.899	2.767	2.500	35,212	298,783	72,555	101,517
C. Manufacturing	20,937	1.171	1.237	18.924	73.662	3.872	2.978	12,355	73,817	60,794	18,468
D. Electricity, gas, steam etc.	925	1.267	2.041	15.340	46.974	3.372	2.926	34,650	321,543	73,488	30,898
E. Water supply, sewerage etc.	1,957	2.129	3.316	10.479	21.034	3.112	2.306	4,353	14,119	59,114	13,886
F. Construction	31,967	1.050	0.738	5.145	14.408	2.298	1.696	2,649	12,075	57,610	14,378
G. Wholesale and retail trade	69,193	1.383	5.722	5.514	15.559	2.518	1.992	6,679	36,576	56,683	21,619
H. Transportation	15,570	1.274	5.125	11.277	50.020	2.794	2.331	7,666	114,439	57,890	25,777
I. Accommodation and food services	15,780	1.239	3.003	3.370	9.242	2.038	1.638	1,488	4,217	48,049	13,443
J. Information and communication	15,495	1.182	3.108	10.968	49.839	2.912	2.604	5,163	29,492	76,131	40,250
L. Real estate	13,050	1.344	2.311	3.541	8.919	2.080	1.728	1,139	4,435	59,727	25,909
M. Knowledge-based services	27,463	1.136	1.231	7.589	30.830	2.753	2.433	2,798	18,008	72,659	47,190
N. Travel agent, cleaning etc.	13,831	1.290	2.322	6.724	19.534	2.668	2.325	3,153	12,084	59,338	36,342
R. Arts, entertainment, recreation	5,804	1.395	2.842	5.765	14.060	2.799	2.420	1,048	22,416	54,942	19,372
S. Other services	13,312	1.126	1.471	4.523	13.985	2.222	1.972	419	2,547	55,563	16,467
All industries	259,195	1.227	3.494	7.414	33.071	2.611	2.223	5,198	49,994	59,311	27,048

Source: Administrative registers, Statistics Denmark. Full population of private sector establishments in Denmark (step 1 in table C.3). 5-digit industry classification based on NACE rev. 2. We exclude the public sector, including the health and education sectors. Revenue and average wage at the firm in 2022 USD.

Table D.4. Establishment characteristics, by industry (estimation sample, all years)

	n. unique estab.	n. estab. per firm		n. of workers per estab.		n. of k-groups per estab.		estab. revenue (1,000 UDS)		average wage (USD)	
commuting zone		mean	st. dev.	mean	st. dev.	mean	st. dev.	mean	st. dev.	mean	st. dev.
A. Agriculture, forestry, and fishery	2,238	1.046	0.581	3.782	5.724	2.372	1.583	2,840	4,262	50,896	11,758
B. Mining and quarrying	134	1.625	2.554	17.891	68.175	3.553	2.662	48,419	357,164	70,750	31,235
C. Manufacturing	8,850	1.171	1.232	24.457	84.925	4.632	2.998	16,004	84,926	61,765	14,128
D. Electricity, gas, steam etc.	306	1.178	1.061	20.179	57.205	4.095	3.001	64,549	456,296	74,386	33,991
E. Water supply, sewerage etc.	438	1.564	1.940	12.983	25.803	3.719	2.517	7,940	19,440	60,679	12,300
F. Construction	8,741	1.059	0.753	7.549	17.836	3.005	1.814	3,949	15,334	59,681	12,588
G. Wholesale and retail trade	21,282	1.356	4.931	7.916	19.278	3.254	2.156	9,794	45,456	59,433	19,340
H. Transportation	4,307	1.322	5.440	16.608	61.499	3.644	2.484	11,221	105,985	59,758	18,708
I. Accommodation and food services	2,018	1.210	1.995	5.749	14.367	3.018	2.008	2,709	6,364	51,705	12,648
J. Information and communication	3,498	1.212	2.826	18.323	63.986	4.175	2.851	8,926	38,975	76,905	24,703
L. Real estate	1,613	1.174	1.121	5.155	12.528	2.922	2.013	2,714	7,538	68,106	28,961
M. Knowledge-based services	6,086	1.160	1.111	12.136	40.652	3.912	2.622	4,899	24,755	72,717	24,334
N. Travel agent, cleaning etc.	2,704	1.182	1.136	7.918	23.035	3.305	2.284	5,492	17,011	62,117	20,404
R. Arts, entertainment, recreation	386	1.085	0.687	9.414	20.773	3.896	2.719	8,397	70,779	60,055	17,245
S. Other services	924	1.144	1.297	7.322	16.345	2.951	2.251	2,260	5,253	57,559	16,769
All industries	63,525	1.219	3.240	11.591	44.041	3.515	2.427	8,909	62,962	61,787	19,573

Source: Administrative registers, Statistics Denmark. Restricted sample of establishments with no missing values for the key estimation variables (step 5 in table C.3). 5-digit industry classification based on NACE rev. 2. We exclude the public sector, including the health and education sectors. Revenue and average wage at the firm in 2022 USD.

TABLE D.5. Labor Supply Parameter Estimates Across k-groups

		IV			OLS		
		β_k	σ_{kg}		β_k	σ_{kg}	
	k-group (k)		CZ 1 (CPH)	Avg. across CZ		CZ 1 (CPH)	Avg. across CZ
1	Female, 26-35, no college	1.701	3.966	3.228	-0.002	5.548	4.342
		[1.386; 2.0786]	[3.014; 4.445]		[-0.020; 0.013]	[4.359; 5.715]	
2	Female, 26-35, college	1.922	5.698	2.803	-0.099	6.352	3.405
		[1.315; 2.5072]	[3.168; 7.324]		[-0.124; -0.072]	[4.627; 6.382]	
3	Male, 26-35, no college	1.392	5.654	3.800	0.321	6.560	4.240
		[1.377; 1.5974]	[4.043; 5.863]		[0.308; 0.325]	[5.179; 6.146]	
4	Male, 26-35, college	2.225	3.926	3.923	0.323	5.057	3.423
		[1.823; 2.6060]	[2.758; 4.612]		[0.305; 0.341]	[3.857; 5.056]	
5	Female, 36-50, no college	1.078	6.169	3.913	0.226	6.347	3.991
		[0.997; 1.3178]	[4.385; 6.516]		[0.216; 0.229]	[4.769; 5.978]	
6	Female, 36-50, college	1.540	4.463	3.776	0.000	4.813	3.657
		[1.234; 2.1316]	[2.946; 5.052]		[0.049 ; 0.079]	[3.565; 4.559]	
7	Male, 36-50, no college	0.874	6.545	3.930	0.272	6.351	4.241
		[0.917; 1.0248]	[4.586; 6.043]		[0.263; 0.274]	[4.680; 5.461]	
8	Male, 36-50, college	1.080	4.403	3.040	0.098	4.530	2.852
		[0.942; 1.3421]	[3.197; 4.265]		[0.087; 0.106]	[3.442; 4.122]	
9	Female, 51-60, no college	1.073	7.524	6.072	0.282	7.353	4.756
		[0.802; 1.3607]	[5.580; 9.417]		[0.275; 0.292]	[6.146; 7.847]	
10	Female, 51-60, college	1.040	6.528	4.727	0.225	5.363	3.170
		[0.663; 1.4159]	[4.538; 9.825]		[0.209; 0.246]	[3.853; 6.203]	
11	Male, 51-60, no college	0.737	7.415	5.622	0.271	7.611	4.765
		[0.723; 0.9123]	[5.105; 7.546]		[0.258; 0.272]	[5.827; 7.074]	
12	Male, 51-60, college	0.938	4.051	3.852	0.157	4.581	3.237
		[0.716; 1.2335]	[2.867; 4.253]		[0.150; 0.170]	[3.460; 4.557]	

Parameter estimates for equation 5.2, OLS and IV. We estimate the parameters separately by k-group. The first column are the point estimates for β_k. The second column shows estimates for the σ_{kg} for the Copenhagen metro area). The third column shows the average σ_{kg} estimate across commuting zones. Bootstrapped 95% confidence intervals in square brackets (Hall, 1992).

Source: Administrative registers, Statistics Denmark.

Table D.6. Labor Supply Elasticities and Markdowns, by k-group

	k-group	IV		OLS	
		Elasticity	Markdown	Elasticity	Markdown
1	Female, 26-35, no college	6.221	0.857	-0.010	-0.010
2	Female, 26-35, college	9.061	0.889	-0.489	-1.144
3	Male, 26-35, no college	6.606	0.858	1.724	0.619
4	Male, 26-35, college	10.747	0.900	1.535	0.591
5	Female, 36-50, no college	5.096	0.824	1.121	0.519
6	Female, 36-50, college	6.141	0.849	0.249	0.197
7	Male, 36-50, no college	4.325	0.800	1.392	0.574
8	Male, 36-50, college	4.100	0.793	0.369	0.265
9	Female, 51-60, no college	8.426	0.871	1.695	0.616
10	Female, 51-60, college	5.755	0.837	0.956	0.479
11	Male, 51-60, no college	4.508	0.788	1.561	0.598
12	Male, 51-60, college	4.070	0.787	0.657	0.388
	Overall	5.790	0.829	1.109	0.331

Estimated labor supply elasticities (eq. 3.1) and markdowns $\left(\mathrm{md}_{kj} = \frac{\mathcal{E}_{kj}}{1+\mathcal{E}_{kj}}\right)$ from the labor supply model. Mean of the pooled (over time) distribution of establishment-level labor supply elasticities and markdowns for each k-group. We estimate the parameters separately by k-group. The first two columns report the IV estimates, the third and fourth columns report the OLS estimates.

Table D.7. Substitution Parameter Estimates Across k-groups

		IV			IV	OLS
	k-group	$\rho_k - 1$	$\delta(\rho_k - 1)$	δ	ρ_k	ρ_k
1	Female, 26-35, no college	0.005	0.005	0.806	1.005	0.985
		[-0.004; 0.012]	[-0.002; 0.010]	[0.804; 0.809]	[0.997; 1.012]	[0.982; 0.988]
2	Female, 26-35, college	0.029	0.028		1.029	0.985
		[0.019; 0.038]	[0.019; 0.037]		[1.019; 1.038]	[0.981; 0.988]
3	Male, 26-35, no college	0.007	0.006		1.007	0.987
		[0.000; 0.014]	[0.000; 0.012]		[1.000; 1.014]	[0.985; 0.989]
4	Male, 26-35, college	0.028	0.029		1.028	0.981
		[0.016; 0.036]	[0.017; 0.037]		[1.016; 1.036]	[0.978; 0.984]
5	Female, 36-50, no college	0.016	0.016		1.016	0.978
		[0.006; 0.026]	[0.007; 0.025]		[1.006; 1.026]	[0.976; 0.980]
6	Female, 36-50, college	0.002	-0.004		1.002	0.992
		[-0.0114; 0.0201]	[-0.018; 0.012]		[0.989; 1.020]	[0.987; 0.996]
7	Male, 36-50, no college	-0.024	-0.022		0.976	0.977
		[-0.033; -0.015]	[-0.030; -0.013]		[0.967; 0.985]	[0.975; 0.979]
8	Male, 36-50, college	-0.065	-0.067		0.935	0.999
		[-0.0832; -0.0505]	[-0.084; -0.053]		[0.917; 0.949]	[0.995; 1.003]
9	Female, 51-60, no college	0.003	0.002		1.003	0.990
		[-0.0094; 0.0159]	[-0.010; 0.013]		[0.991; 1.016]	[0.987; 0.993]
10	Female, 51-60, college	-0.027	-0.034		0.973	1.017
		[-0.0538; 0.0022]	[-0.060; -0.004]		[0.946; 1.002]	[1.009; 1.026]
11	Male, 51-60, no college	-0.016	-0.013		0.984	0.985
		[-0.0276; -0.0053]	[-0.025; -0.003]		[0.972; 0.995]	[0.981; 0.988]
12	Male, 51-60, college	-0.036	-0.041		0.964	1.026
		[-0.053; -0.007]	[-0.058; -0.014]		[0.948; 0.993]	[1.020; 1.035]

Parameter estimates for the production function, IV. The first two columns are the point estimates for $(\rho_k - 1)$ and $\delta(\rho_k - 1)$ from equation 5.6. The third and fourth columns show the implied values for δ and ρ_k. The fifth column shows the OLS estimate for ρ_k. Bootstrapped 95% confidence intervals in square brackets.

Source: Administrative registers, Statistics Denmark.

Table D.8. Distribution of Labor Demand Elasticities η_{kjt}, by k-group.

		η_{kjt}			
	k-group	Mean	Median	P10	P90
1	Female, 26-35, no college	-27.070	-9.859	-70.952	-2.187
2	Female, 26-35, college	21.871	-8.528	-83.959	102.098
3	Male, 26-35, no college	9.423	-5.557	-23.556	-1.887
4	Male, 26-35, college	-60.934	-9.597	-75.196	71.058
5	Female, 36-50, no college	-9.958	-7.228	-37.837	-2.001
6	Female, 36-50, college	-28.406	-12.042	-52.689	-2.990
7	Male, 36-50, no college	-4.003	-2.961	-7.104	-1.488
8	Male, 36-50, college	-4.884	-4.326	-8.573	-2.058
9	Female, 51-60, no college	-24.150	-10.801	-52.521	-2.658
10	Female, 51-60, college	-13.663	-12.035	-27.345	-2.845
11	Male, 51-60, no college	-6.225	-4.537	-12.166	-1.964
12	Male, 51-60, college	-8.461	-7.265	-16.165	-2.640

Moments of the firm-level labor demand elasticities η_{kjt} as defined in Section 3.2, eq. 3.5.

Table D.9. Morishima Elasticity of Substitution Between k-groups.

kgroup (k)		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Female, 26-35, no college	**1**	0	-42	-161	-214	-78	510	53	199	-188	2305	276	259
Female, 26-35, college	**2**	-168	0	-116	-36	-69	-739	-30	25	-138	38	-25	14
Male, 26-35, no college	**3**	-183	-45	0	-38	-62	-471	23	-48	-189	12	32	27
Male, 26-35, college	**4**	135	-35	-123	0	-63	778	37	16	204	673	170	87
Female, 36-50, no college	**5**	-156	-34	-130	-35	0	-446	-14	19	-144	26	-13	24
Female, 36-50, college	**6**	-625	2	-95	-347	20	0	-231	25	-470	3160	239	365
Male, 36-50, no college	**7**	54	-48	-93	8	-88	304	0	17	156	14	60	23
Male, 36-50, college	**8**	192	-27	-80	-3	-59	690	43	0	285	178	97	34
Female, 51-60, no college	**9**	-335	-64	-206	-284	-92	55	199	239	0	2411	313	243
Female, 51-60, college	**10**	727	-32	-290	106	-73	2681	41	-16	594	0	110	17
Male, 51-60, no college	**11**	185	-42	-131	29	-121	430	51	15	173	-129	0	16
Male, 51-60, college	**12**	388	-46	-143	42	-69	1609	41	6	222	107	78	0

Each cell is the mean Morishima elasticity of substitution calculated across all firms which employ both types of labor.

Table D.10. Variance Decomposition of Counterfactual Wages

Scenario 1

Counterfactual Exercise:	Truth	A	B	C	D	E
Variance of Log Wages	0.1285	0.427	0.3346	0.2764	0.2031	0.0005
Variance of Log Markdown	0.0067	0.0067	0.0004	0.0004	0.0003	0.0003
Variance of Log MRPL	0.1394	0.3967	0.3309	0.2723	0.1987	0.0001
2 × Covariance	-0.0176	0.0237	0.0033	0.0037	0.0041	0.0001

Scenario 2

Counterfactual Exercise:	Truth	C	A	B	D	E
Variance of Log Wages	0.1285	0.086	0.3876	0.2764	0.2031	0.0005
Variance of Log Markdown	0.0067	0.0067	0.0067	0.0004	0.0003	0.0003
Variance of Log MRPL	0.1394	0.08	0.3349	0.2723	0.1987	0.0001
2 × Covariance	-0.0176	-0.0007	0.046	0.0037	0.0041	0.0001

Scenario 3

Counterfactual Exercise:	Truth	C	D	A	B	E
Variance of Log Wages	0.1285	0.086	0.0912	0.2827	0.2031	0.0005
Variance of Log Markdown	0.0067	0.0067	0.0067	0.0066	0.0003	0.0003
Variance of Log MRPL	0.1394	0.08	0.0942	0.2405	0.1987	0.0001
2 × Covariance	-0.0176	-0.0007	-0.0097	0.0356	0.0041	0.0001

Scenario 4

Counterfactual Exercise:	Truth	A	B	D	C	E
Variance of Log Wages	0.1285	0.427	0.3346	0.3087	0.2031	0.0005
Variance of Log Markdown	0.0067	0.0067	0.0004	0.0004	0.0003	0.0003
Variance of Log MRPL	0.1394	0.3967	0.3309	0.3048	0.1987	0.0001
2 × Covariance	-0.0176	0.0237	0.0033	0.0035	0.0041	0.0001

Scenario 5

Counterfactual Exercise:	Truth	D	C	A	B	E
Variance of Log Wages	0.1285	0.1908	0.0912	0.2827	0.2031	0.0005
Variance of Log Markdown	0.0067	0.0067	0.0067	0.0066	0.0003	0.0003
Variance of Log MRPL	0.1394	0.21	0.0942	0.2405	0.1987	0.0001
2 × Covariance	-0.0176	-0.0259	-0.0097	0.0356	0.0041	0.0001

Scenario 6

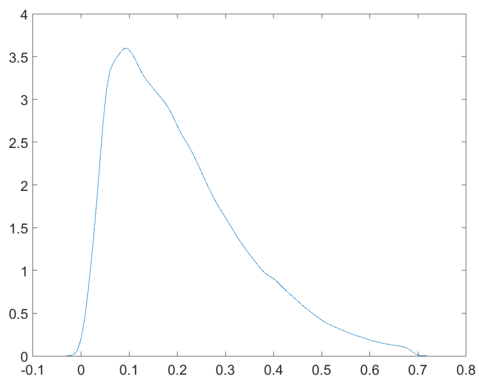
Counterfactual Exercise:	Truth	B	A	D	C	E
Variance of Log Wages	0.1285	0.1573	0.3346	0.3087	0.2031	0.0005
Variance of Log Markdown	0.0067	0.0004	0.0004	0.0004	0.0003	0.0003
Variance of Log MRPL	0.1394	0.157	0.3309	0.3048	0.1987	0.0001
2 × Covariance	-0.0176	0	0.0033	0.0035	0.0041	0.0001

Scenario 7

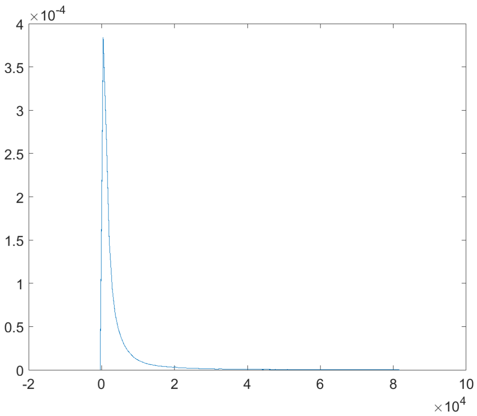
Counterfactual Exercise:	Truth	D	C	B	A	E
Variance of Log Wages	0.1285	0.1908	0.0912	0.1101	0.2031	0.0005
Variance of Log Markdown	0.0067	0.0067	0.0067	0.0004	0.0003	0.0003
Variance of Log MRPL	0.1394	0.21	0.0942	0.1102	0.1987	0.0001
2 × Covariance	-0.0176	-0.0259	-0.0097	-0.0005	0.0041	0.0001

Counterfactual estimates of the variance of log wages, decomposed into the variances of log markdowns and log MRPL and (2×) the covariance from eq.(5.4), for 7 different decomposition scenarios. In each scenario, each column represents a cumulative counterfactual exercise, where the effect is inclusive of previous columns. For example, Scenario 1 column 3 includes both exercise A and B and Column 4 includes exercises A, B and C. Exercise A sets $u_{jk} = \overline{u}$, B sets $\beta_k = \overline{\beta}$ and $\sigma_{gk} = \overline{\sigma}$, C sets $\gamma_{kj} = \overline{\gamma}$ and $\rho_k = \overline{\rho}$, D sets $\theta_j^{\alpha_j} = \overline{\theta^{\alpha}}$ and $\alpha_j = \overline{\alpha}$, and E sets $\alpha_j = 1$. The overline represents the observation-weighted mean, except in D where it is the median.

FIGURE D.1. Distribution of Scale (α_{jt}) and Firm Productivity ($\tilde{\theta}_{jt}^{\alpha_{jt}}$).



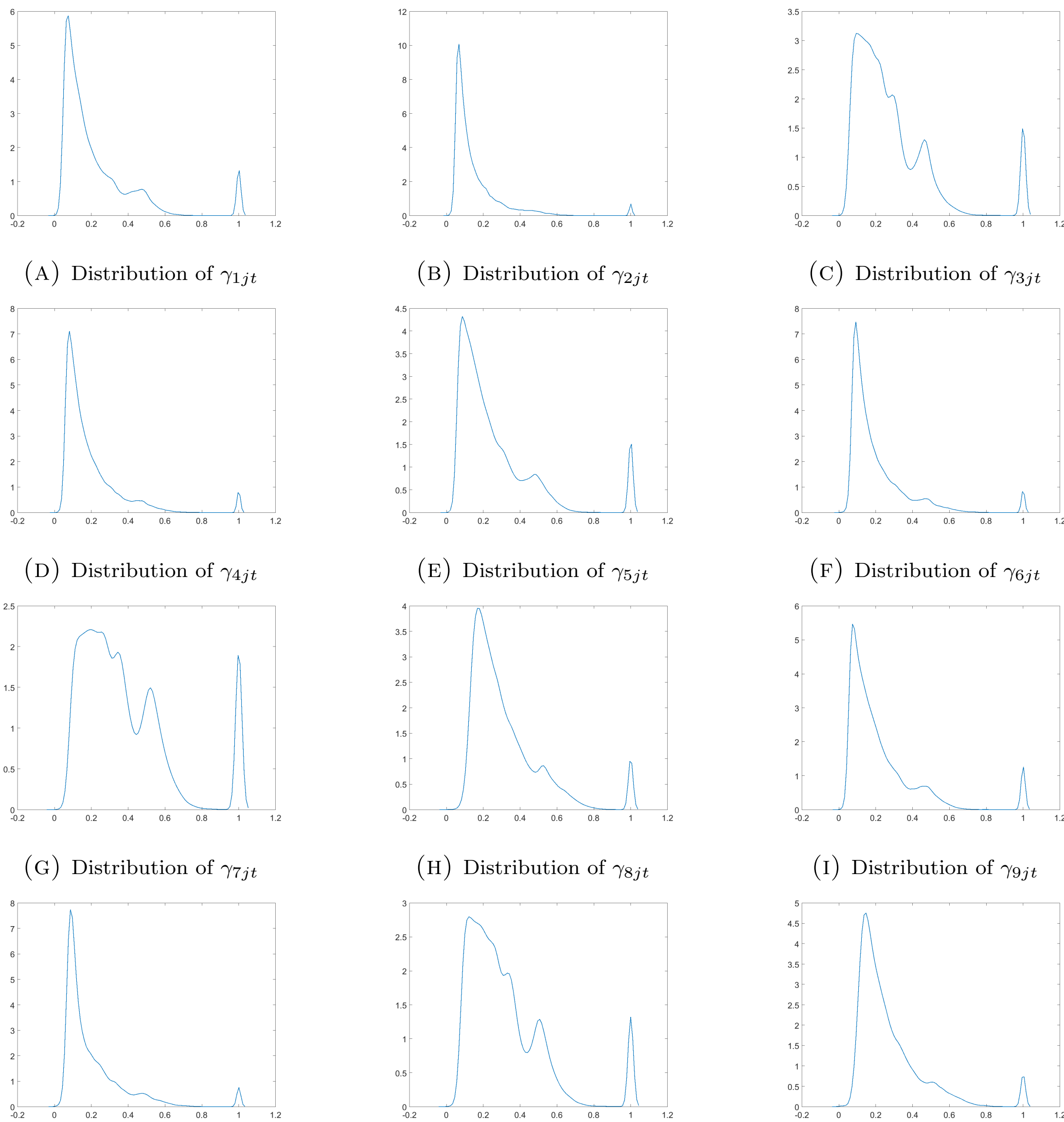
(A) Distribution of α_{jt}



(B) Distribution of $\tilde{\theta}_{jt}^{\alpha_{jt}}$

Panel (a) shows the distribution of the scale parameter α_{jt} (eq. 5.8). The mean of this distribution is 0.214 and the median is 0.181. Panel (b) shows the distribution of productivity term $\tilde{\theta}_{jt}^{\alpha_{jt}}$, truncated at the 99th percentile (eq. 5.9). The mean of the truncated distribution is $6,693$ (in 2021 Danish krona). The 90-10 ratio for $\tilde{\theta}_{jt}^{\alpha_{jt}}$ taken over all private sector firms in the economy is 24.3.

FIGURE D.2. Distribution of Normalized Labor Productivity (γ_{kjt}) for each k-group.



(A) Distribution of γ_{1jt} (B) Distribution of γ_{2jt} (C) Distribution of γ_{3jt}

(D) Distribution of γ_{4jt} (E) Distribution of γ_{5jt} (F) Distribution of γ_{6jt}

(G) Distribution of γ_{7jt} (H) Distribution of γ_{8jt} (I) Distribution of γ_{9jt}

(J) Distribution of γ_{10jt} (K) Distribution of γ_{11jt} (L) Distribution of γ_{12jt}

The 12 panels show the distribution of the normalized productivity parameter γ_{kjt} for each of the 12 k-groups (eq. 5.7). The mean and medians of these distributions by k-group are in Table D.8.

References

Azar, José A, Steven T Berry, and Ioana Marinescu, "Estimating labor market power," Technical Report, National Bureau of Economic Research 2022.

Berry, Steven, James Levinsohn, and Ariel Pakes, "Automobile prices in market equilibrium," *Econometrica*, 1995, pp. 841–890.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa, "A distributional framework for matched employer employee data," *Econometrica*, 2019, *87* (3), 699–739.

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

Carlson, David and Thomas L Markham, "Schur complements of diagonally dominant matrices," *Czechoslovak Mathematical Journal*, 1979, *29* (2), 246–251.

Eckert, Fabian, Mads Hejlesen, and Conor Walsh, "The return to big-city experience: Evidence from refugees in Denmark," *Journal of Urban Economics*, 2022, *130*, 103454.

Fiedler, Miroslav and Vlastimil Pták, "On matrices with non-positive off-diagonal elements and positive principal minors," *Czechoslovak Mathematical Journal*, 1962, *12* (3), 382–400.

Frommer, Andreas, "Generalized nonlinear diagonal dominance and applications to asynchronous iterative methods," *Journal of Computational and Applied Mathematics*, 1991, *38* (1-3), 105–124.

Gan, Tai-Bin, Ting-Zhu Huang, and Jian Gao, "A note on generalized nonlinear diagonal dominance," *Journal of mathematical analysis and applications*, 2006, *313* (2), 581–586.

Garin, Andrew and Filipe Silvério, "How Responsive Are Wages to Firm-Specific Changes in Labour Demand? Evidence from Idiosyncratic Export Demand Shocks," *The Review of Economic Studies*, 2023.

Hagedorn, Marcus, Tzuo Hann Law, and Iourii Manovskii, "Identifying equilibrium models of labor market sorting," *Econometrica*, 2017, *85* (1), 29–65.

Hall, Peter, "On bootstrap confidence intervals in nonparametric regression," *The Annals of Statistics*, 1992, pp. 695–711.

Hazell, Jonathon, Christina Patterson, Heather Sarsons, and Bledi Taska, "National wage setting," Technical Report, National Bureau of Economic Research 2022.

Hummels, David, Rasmus Jørgensen, Jakob Munch, and Chong Xiang, "The wage effects of offshoring: Evidence from Danish matched worker-firm data," *American Economic Review*, 2014, *104* (6), 1597–1629.

Kroft, Kory, Yao Luo, Magne Mogstad, and Bradley Setzler, "Imperfect competition and rents in labor and product markets: The case of the construction industry," Technical Report, National Bureau of Economic Research 2023.

Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler, "Imperfect competition, compensating differentials, and rent sharing in the US labor market," *American Economic Review*, 2022, *112* (1), 169–212.