AN EMPIRICAL FRAMEWORK FOR MATCHING WITH IMPERFECT COMPETITION

MONS CHAN, KORY KROFT, AND ISMAEL MOURIFIÉ

The University of Toronto, The University of Toronto & NBER, and the University of Toronto

Abstract. This paper considers a static, many-to-one matching model of the labor market. Firms face inelastic labor supply curves and hence charge an endogenous firm-specific markdown below marginal product. We assume that firms operate in an oligopsony labor market and thus allow for strategic interactions in wage setting. We provide a tractable characterization of the equilibrium and demonstrate existence and uniqueness. This characterization of the model equilibrium allows us to derive a rich set of comparative statics and then to gauge the relative contributions of worker skill, preference for amenities and imperfect competition on equilibrium wage inequality. We also illustrate identification of structural parameters and the estimation of complementarity between workers and firms using matched employer-employee data on the population of Danish workers.

PRELIMINARY PLEASE DO NOT CIRCULATE

Keywords: Compensating differentials, Inequality, Imperfect competition, Matching equilibrium, *kmeans* clustering, Oligopsony, Strategic interactions, Sorting.
JEL subject classification: C78, D3, J21, J23, J31, J41.

1. Introduction

A large literature in labor economics has sought to understand the determinants of earnings inequality. Among the key explanations are that workers have different skills (the Roy model and

Date: January 31, 2020. We thank Victor Aguirregabiria, Karen Bernhardt-Walther, Antoine Djogbenou, Yao Luo, Suresh Naidu, Marcin Peski, and Aloysius Siow for helpful discussions. Correspondence address: Department of Economics, University of Toronto, 150 St. George Street, Toronto ON M5S 3G7, Canada. Emails: mons.chan@utoronto.ca, kory.kroft@utoronto.ca, ismael.mourifie@utoronto.ca.

human capital), that similarly skilled workers make different choices based on their preferences for working at firms that vary in their offered amenities (compensating differentials), and that the market treats observationally similar people differently (imperfect competition). Much of the literature seeks to test these theories in isolation, although there are some recent exceptions (see, for example, Postel-Vinay and Robin (2002) and Taber and Vejlin (2016)). This paper develops a static, many-to-one matching model of the labor market that combines these different aspects of wage determination into a unifying framework.

On one side of the market, there are a finite number of firms with heterogeneous production functions which depend on workers' skills. Workers differ in skill dimensions that are allowed to be match-specific, meaning that the same employee might be more productive working for a specific firm compared to other firms. Allowing for match-specific skills enables our model to capture key features of the Roy model. On the other side of the market are a large number of heterogeneous individuals that have random preferences over wages and employer amenities (both deterministic and stochastic), similar to Rosen (1974, 1976). We assume that each individual is characterized by a "type" which is observable to firms, but not necessarily to the econometrician. Each individual chooses a firm (or unemployment) to maximize utility. The heterogeneity in non-pecuniary preferences along the horizontal and vertical dimensions across workers implies that the firm-specific choice probabilities are characterized by a finite wage elasticity. Firms may be able to pay their workers below their marginal products in equilibrium (without fear of losing them) since workers may value their non-pecuniary amenities.[1]

The presence of upward-sloping *firm-specific* labor supply curves gives employers market power to set wages.[2] We assume that firms post wages unilaterally for each worker type in order to minimize the cost of producing a given level of output before the worker and employer meet, taking into

[1]One difference between our framework and that of Rosen (1986) is we take amenities to be exogenous and the only choice variable for firms is wages. On the other hand, Rosen assumes that wages are exogenous from the firm's perspective, and firms choose the value of the amenities. Additionally, in Rosen the preference for the amenity does not vary across firms, unlike our setting where it does.

[2]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.

account both labor supply and also accounting for wages set by other firms in the market.[3] Thus, our model features an *oligopsony* market structure. We derive the optimal wage for a worker of a given type and provide a characterization for the equilibrium wage and distributional matching functions (hereafter referred to as the EWF & DMF, respectively).

We demonstrate two key sets of results. First, we show that when the stochastic part of individuals' non-pecuniary preferences are i.i.d. random variables distributed according to the Extreme Value Type-I (Gumbel) distribution – what we call "the Logit economy" – and workers are perfect substitutes in production, the found an analytical form for the EWFs and DMFs that depend on (i) the workers match-specific productivity, (ii) the deterministic part of workers non-pecuniary preferences for firm amenities, and (iii) the share of unemployed of each worker type. The first two terms allow us to investigate how changes in workers match-productivity distribution and non-pecuniary preferences affect the equilibrium wage, while the vector of the unemployed share can be viewed as a "sufficient statistic" that captures the impact of strategic interactions across employees on the equilibrium wage; it is shown to be the unique solution of K non-linear equations with K unknowns, where K represents the total number of worker types in the economy. This tractable characterization of the model equilibrium allows us to derive a rich set of comparative statics that are useful to gauge the relative contributions of worker skill, preference for amenities and imperfect competition on equilibrium wage inequality.

Our second set of results are related to identification. We innovate by making use of both the EWFs and DMFs to recover key structural parameters of the model. Since Choo and Siow (2006), identification of structural parameters in the matching literature has mainly used *only* the DMF.[4] On the other hand, in the empirical labor literature, the focus has been mainly the EWF. Here, the model we entertain features a tractable mapping between both the EWF and DMF and the structural parameters. Therefore, both contain useful information for identification and we show how one may exploit this insight to estimate the structural parameters. One of our main objectives is the identification of a measure of complementarity between workers and firms, namely the *Aggregate Local Complementarity* (ALC) (see Graham, 2011).[5] Testing for complementarities has been an important topic of interest in the labor literature. We show that when the firm sets wages based only on characteristics observed by the econometrician, the ALC can be identified on a single cross-section of data, using both the EWF and the DMF. However, when the firm sets wages based

[3]In our model, employers do not have full information over workers preferences since they cannot observe the idiosyncratic part of individual's preferences. Thus, firms cannot perfectly wage discriminate and extract all workers surplus.

[4]This is mainly motivated by the fact that in the marriage literature the transfers (wages in our case) are not observed.

[5]The ALC measures the aggregate difference between the incremental productivity return associated with hiring a type k versus k' worker across firm j versus j'.

on characteristics that are unobserved to the econometrician (point) identification is not possible without additional information. In such a context we propose a two-step procedure that requires the use of panel data. The first step uses the Bonhomme, Lamadon and Manresa (2017)'s *kmeans* approach to recover worker unobserved heterogeneity which is finite in our framework. Once the unobserved types are recovered, the second-step consistently estimates the ALC as if the unobserved types were known by the econometrician.

Our paper relates to and builds on several strands of the literature. First, 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) allows workers to be complements in the firms production functions and peer effects in an ex-post Nash-bargaining model. It is also noteworthy that these papers and most of their extensions do not consider unobserved heterogeneity and more importantly consider a competitive equilibrium market with perfect information.[6] Our contribution to this literature is to consider a wage-posting model with worker-level unobserved heterogeneity and imperfect information. None of the existence results in the matching literature directly apply to our context. We show that in the Logit economy, when workers are perfectly substitutable, the equilibrium distribution of matches associated with our many-to-one TU model is equivalent to solve K distinct non-linear fixed-point problems. We show that a solution for those equations always exists and is unique. A related paper is Azevedo (2014) who considers an imperfectly competitive many-to-one matching market. However, he does not consider unobserved heterogeneity and he mostly focuses on the case where firms compete on quantities given exogenously fixed wages.

Second, our paper relates to a growing literature that seeks to understand whether there are microfoundations that can rationalize the log wage equation which is additively separable in a worker- and firm-fixed effect in Abowd, Kramarz and Margolis (1999) (hereafter "AKM"). Most closely related to this paper are Card et al. (2018) and Lamadon, Mogstad and Setzler (2018) who propose empirical structural models with imperfect competition. These papers adopt an "atomistic" firm assumption whereby firms do not internalize the impact of their wage setting on the market wage. This considerably simplifies the problem by eliminating strategic interactions.[7] Thus, these papers consider models of *monopsonistic competition*. Our paper builds on these papers by considering a

[6]Rare exceptions which consider unobserved heterogeneity but assuming a competitive equilibrium market with perfect information are Dupuy et al (2017, 2018).

[7]The importance of having an empirical framework with strategic interactions has been emphasized by Card et al (2018, section F).

model of *oligopsony*. This relates to recent work by Berger, Herkenhoff and Mongey (2019) who also develop a model of oligopsony and use it to study the change in concentration over time. Finally, although our paper does not incorporate dynamic considerations, it relates to the search-and-matching literature which incorporate firm and worker heterogeneity.[8] Much of this literature focuses on the extent to which the correlation between worker and firm fixed effects in the AKM wage equation reveals information about the nature of sorting between workers and firms. Our paper is most closely related to Taber and Vejlin (2016) 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.[9]

Third, our paper relates to the classic model in Roy (1951) and subsequent research. Roy proposed a model of earnings with sorting on sector-specific skills, and conjectured that favourable sorting reduces earnings inequality. Using respectively, a parametric and nonparametric framework, Heckman and Honoré (1990) and Mourifié et al (2018) study the validity of this argument in a partial equilibrium model where the price of skills are exogenously determined. Our contribution is to study the implications of sorting for equilibrium wage inequality in a general equilibrium model where employers have the ability to set wages, and employees have non-pecuniary preferences for employer amenities.

Lastly our paper relates to the literature on compensating differentials. This literatures considers models that typically assume perfect competition and show that the equilibrium wage gap between two equally productive workers is driven by their different valuation of firm amenities. Empirically this gap is therefore used to price non-income job characteristics. Our contribution is to understand the extent to which the predictions of competitive models survive in an imperfect competition environment.[10] Another distinguishing feature of our model is that it relaxes the restriction in other papers (e.g., Card et al (2018) and Lamadon, Mogstad and Setzler (2018)) that two equally productive workers at a firm have the same equilibrium wage –even if they have different preferences for the firm's amenities. A paper that is related to ours is Sorkin (2018) who considers a search model and shows how voluntary flows between firms can reveal the extent to which compensating differentials matter. We follow Sorkin by using information on matching between workers and firms to learn about preferences for non-pecuniary benefits. This is closer in spirit to the traditional IO approach which uses product-market shares to identify brand preferences. Moreover, while Sorkin

[8]Key papers are Eeckhout and Kircher (2011), Lopes de Melo (2013), Lentz (2010), Lise, Meghir and Robin (2016), Bagger and Lentz (2016), Taber and Vejlin (2016) and Hagedorn, Law and Manovski (2017).

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

[10]See for instance the discussion in Brown (1980, section V).

does not consider worker-level heterogeneity or unemployment, our work incorporates these features explicitly.

The outline of the rest of the paper is as follows: Section 2 presents our theoretical and empirical framework. Section 3 proposes a tractable characterization of the equilibrium wage and distributional matching function and discuss the existence and uniqueness of the equilibrium. Section 5 concerns identification and estimations of key parameters. Section 6 proposes an empirical illustration of our approach (in progress). The last section concludes. Proofs of the main results are collected in the appendix.

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, $M < \infty$. One way to rationalize this is by noting that the *proportion* of individuals in each category, $\mathfrak{m}_k \equiv \frac{m_k}{M}$, in a finite population is consistent with the proportion in an infinite population. More precisely, note that $\frac{m_k}{M}$ remains constant as m_k and $M = \sum_{k\in\mathcal{K}} m_k$ 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] We consider that each type k can be divided in two subgroups of types, i.e. $k \equiv (\bar{k}, \tilde{k})$, where $\bar{k}$ is defined based on the underlying vector of characteristics $\bar{X}$ that are observed both by the econometrician and firms (i.e. gender, education, age, immigration status, etc) while $\tilde{k}$ is defined based on the set of characteristics $\tilde{X}$ that are observable only to firms but not to the econometrician (i.e. ability, non-cognitive skill, etc). 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 are differentiated by their unobservable (both to firms and the econometrician)

[11]First, in a finite population there is almost always a profitable deviation which may complicate the analysis of the existence of a stable equilibrium. Second, below we derive choice probabilities, in order to connect these choice probabilities to market shares, we require there to be a continuum of individuals in the population.

[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 $K = K_1 \times ... \times K_{|\mathcal{D}|}$.

characteristics and taste for different firms. Each individual i chooses to work at a firm or to be unemployed, and each firm chooses wages associated with each worker type.

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 $w_{k_i j} \equiv w_{kj} \in [0, \infty)$ to work at firm j be given by:

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

where $\ln u_{kj}$ —such that $u_{kj} \in (0, \infty)$— represents the deterministic non-pecuniary part of the worker potential utility U_{ij} net of the non-pecuniary part of the worker utility of being unemployed, and ϵ_{ij} denotes the error term (idiosyncratic payoff) which is unknown to firms.[13] Individual i's utility of being unemployed is given by:

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

where $w_{k0} \in (0, \infty)$ are the non-employment wage (e.g., unemployment benefit level) and the non-pecuniary benefit of non-employment, respectively.[14]

Notice that in this framework, a type k workers take wages as given and have no market power over firms. Special cases of the ARUM specification in eq (2.1) have been proposed in the literature. Card et al (2018) consider the case where $u_{kj} = u_{\bar{k}j}$ and $w_{kj} = w_{\bar{k}j}$ meaning that the econometrician has the same information as firms. Lamadon, Mogstad and Setzler (2018) consider the case where $\tilde{k} = (\tilde{k}_1, \tilde{k}_{-1})$ where $\tilde{k}_1$ represents worker productivity and $\tilde{k}_{-1}$ is a vector of characteristics that only affects non-pecuniary preferences, i.e. $w_{kj} = w_{\tilde{k}_1 j}$ and $u_{kj} = u_{\tilde{k}_{-1} j}$, although they allow $\tilde{k}_1$ and $\tilde{k}_{-1}$ to be arbitrarily correlated across individuals. These specifications impose exclusion restrictions that are potentially motivated by the specific questions they are interested in; however our methodology does not require us to impose them.[15] Given the potential wage streams $\{w_{kj}\}_{0 \leq j \leq J}$, individual i chooses according to:

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

where $v_{kj} \equiv \ln u_{kj} + \ln w_{kj}$, $v_{k0} \equiv \ln w_{k0}$. 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*, see McFadden (1978, 1981):

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

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

[13]Notice that not having a parameter β_{kj} in front of $\ln w_{kj}$ in eq (2.1) is not a restriction but only an innocuous normalization since $\ln u_{kj}$ is non-parametrically defined and no specific parametric distribution has been imposed on ϵ_{ij} so far.

[14]Equivalently, the utilities could have been written $U_{ij} = \ln \tilde{u}_{kj} + \ln w_{kj} + \epsilon_{ij}$, for $j \in \mathcal{J} \cup \{0\}$. However, since $\tilde{u}_{kj}$, and $\tilde{u}_{k0}$ cannot be separately identified, we directly use $u_{kj} = \frac{\tilde{u}_{kj}}{\tilde{u}_{k0}}$.

[15]Notice that Card et al (2018) and Lamadon, Mogstad and Setzler (2018) do not model employment, i.e. eq (2.2).

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[16]

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

and therefore, the labor supply function is given by:

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

where $(\mu_{kj})^s$ represents the number of type k workers that would like to work for 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 depend on the commonly used Logit framework. Later we will analyze the special cases of Logit in section 3.2, the Nested Logit in Appendix A.3 and will discuss their implications on the equilibrium characterization.

Firms: Wage-Posting framework. Each firm j has a production function given by:

$$Y^j = F^j(\mu_{\cdot j}), \tag{2.6}$$

where $\mu_{\cdot j} = (\mu_{1j}, ..., \mu_{Kj})$. For simplicity, we ignore capital and intermediate inputs. Each firm observes the labor supply stream $\{(\mu_{kj})^s\}_{(k,j)\in(\mathcal{K}\times\mathcal{J})}$ and posts a wage offer before meeting a potential worker who then makes a take-it or leave-it decision. Each firm j plays its best response strategy taking other firms' strategies as given. Given knowledge of the labor supply function (2.5), type j's firm best response consists in posting a stream of worker's type-specific wages that minimizes the cost of labor services. More precisely, the firm j best response is obtained as follows:[17]

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

where

$$\mu_{kj} = m_k \frac{\partial G_{k\cdot}(v_{k\cdot})}{\partial v_{kj}}, \ \ (k,j) \in (\mathcal{K} \times \mathcal{J}); \tag{2.7}$$

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

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

[17]We can equivalently consider this following firm's minimization problem:

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

Assumption 2. *(i) We assume that the minimum acceptable level of output for each firm is positive, i.e.* $Y_j > 0,\ j \in \mathcal{J}$. *(ii) The firms' production functions* $F^j(.); j \in \mathcal{J}$ *are assumed to be (a) continuously differentiable, (b) non-constant and non-decreasing in each of its arguments —isotone function, and have zero production with zero labor inputs, i.e.* $F^j(0) = 0$.

Under Assumption 2 (ii-a) the Karush-Kuhn-Tucker (KKT) necessary conditions for optimality of the firm's optimization problem are given by:[18]

(A-1) $\mu_{kj} + w_{kj}\frac{\partial \mu_{kj}}{\partial w_{kj}} - \lambda_j \frac{\partial \mu_{kj}}{\partial w_{kj}} F^j_k(\mu_{\cdot j}) \geq 0$,
(A-2) $w_{kj} \geq 0$,
(A-3) $w_{kj}\left[\mu_{kj} + w_{kj}\frac{\partial \mu_{kj}}{\partial w_{kj}} - \lambda_j \frac{\partial \mu_{kj}}{\partial w_{kj}} F^j_k(\mu_{\cdot j})\right] = 0$,
(A-4) $F^j(\mu_{\cdot j}) - Y_j \geq 0$,
(A-5) $\lambda_j \geq 0$,
(A-6) $\lambda_j \left[F^j(\mu_{\cdot j}) - Y_j\right] = 0$, 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 $\mu_{kj} = 0$. Under Assumptions 2 (i)-(ii-b), (A-4) is not violated if there exist k such $\mu_{kj} > 0$ 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 $s_{kj} \equiv \frac{\mu_{kj}}{m_k}$ denote the share of type k workers working for firm j, and 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 $\mu_{kj} > 0$ for all $k \in \mathcal{C}^j$. We then have

$$w_{kj} = \lambda_j F^j_k(\mu_{\cdot j}) \frac{\mathcal{E}_{kj}}{1 + \mathcal{E}_{kj}}, \text{ for all } k \in \mathcal{C}^j \tag{2.8}$$

or equivalently,

$$\frac{w_{kj}\left(\frac{1+\mathcal{E}_{kj}}{\mathcal{E}_{kj}}\right)}{F^j_k(\mu_{\cdot j})} = \lambda_j, \quad \text{for all } k \in \mathcal{C}^j, j \in \mathcal{J}. \tag{2.9}$$

where $\mathcal{E}_{kj} \equiv \frac{w_{kj}}{\mu_{kj}} \frac{\partial \mu_{kj}}{\partial w_{kj}} \equiv \frac{w_{kj}}{s_{kj}} \frac{\partial s_{kj}}{\partial w_{kj}} > 0$ is the elasticity of labor supply at the optimal wage, $\lambda_j > 0$ is the Lagrange multiplier that represents the marginal cost of production that the firm will equate to marginal revenue at the optimal choice of Y^j, and $F^j_k(\mu_{\cdot j}) \equiv \frac{\partial F^j(\mu_{\cdot j})}{\partial \mu_{kj}} > 0$. Using the latter formulation, we see that firm j best response equates the marginal cost of output across all worker types $k \in \mathcal{C}^j$. For $k \notin \mathcal{C}^j$, firm j's best response is the corner solution $w_{kj} = 0$. 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. Recall that under Assumption 1, the social surplus function

[18]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.

is known to be convex,[19] then $\mathcal{E}_{kj} \geq 0$, and also Assumption 2 (ii-b) ensures that $F_k^j(\mu_{\cdot j}) \geq 0$. Furthermore, we can write the elasticity in terms of the social surplus function as follows:

$$\mathcal{E}_{kj} = \frac{m_k}{\mu_{kj}} \frac{\partial^2 G_{k\cdot}(v_{k\cdot})}{\partial^2 v_{kj}} = \frac{1}{s_{kj}} \frac{\partial^2 G_{k\cdot}(v_{k\cdot})}{\partial^2 v_{kj}}, \quad s_{kj} > 0.$$

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

$$w_{kj} = \lambda_j F_k^j(\mu_{\cdot j}) \frac{\frac{\partial^2 G_{k\cdot}(v_{k\cdot})}{\partial^2 v_{kj}}}{s_{kj} + \frac{\partial^2 G_{k\cdot}(v_{k\cdot})}{\partial^2 v_{kj}}} 1\{s_{kj} > 0\} \tag{2.10}$$

Notice that in such a case all firms are in Nash equilibrium. So far we have described the behaviour of each side of the market. Now, we will 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 and given firm' production functions, an equilibrium outcome (s, w) consists of a distributional worker-firm matching function and an equilibrium wage equation and 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, and the following population constraint holds*

$$\sum_{j \in \mathcal{J}} s_{kj} + s_{k0} = 1, \quad k \in \mathcal{K}. \tag{2.11}$$

Proposition 1. *Under Assumptions 1, and 2, the equilibrium outcome must satisfies equations (2.7), (2.10) and (2.11).*

Remark 1 (Exogenous vs endogenous amenities)**.** *In this section, we have considered a case with exogenous amenities, i.e. u_{kj}, more precisely firms do not have power to set the level of amenities. In Appendix B, we consider the case where firms have power to set both the wages and also the level of amenities. As will be seen, our main approach to provide a tractable characterization of the equilibrium applies also to the endogenous amenities case. However, for sake of simplicity, in the main text we will focus on the exogenous amenities case. For endogenous amenities case, please see Appendix B.*

Remark 2. *In this section, and in Definition 1, we consider that labor supply is derived from an additive random utility maximization, the ARUM. However, workers observed choices may not always be rationalized by an ARUM but can be rationalized by either a Non-Additive RUM or alternative theory of choices like the elimination by aspects discussed in Tversky (1972). In such a context, this section and Definition 1 has to be adapted accordingly. However, the structure of the model remains the same.*

[19] See McFadden (1981) or Shi et al (2018, Lemma 2.1)

3. EQUILIBRIUM WAGE AND DISTRIBUTIONAL MATCHING FUNCTIONS

In this section, we specialize our structural model by specifying a particular distribution for the error terms and specifying the production function. This allows us to solve for the equilibrium wage and distributional matching functions and examine the role of primitives. We begin by introducing the following assumption on the production function:

Assumption 3 (Linear production functions)**.** *The production functions of firm j is:*

$$F^j(\mu_{\cdot j}) = \theta_j\Big(\sum_{k\in\mathcal{K}} \gamma_{kj}\mu_{kj}\Big),$$

where $\theta_j > 0$ is a firm-specific productivity shifter and $\gamma_{kj} > 0$ is the type k worker productivity in firm j.

Assumption 3 imposes that workers are perfect substitutes. While this can certainly be viewed as restrictive, a relaxation of this assumption will be discussed later on. For now, it allows us to make a clear link with the existing literature. However, below we relax this assumption and provide a generalization. Some papers like Postel-Vinay and Robin (2002) and Lamadon, Mogstad and Setzler (2018) consider that skill is one dimensional and does not vary across firms, i.e. $\gamma_{kj} = \gamma_k$. Lamadon, Mogstad and Setzler (2018) further consider that $\gamma_{(\bar{k},\tilde{k})} = \gamma_{\tilde{k}_1}$. Here, we do not impose these restrictions and instead follow Roy (1951) and more recently Taber and Vejlin (2016) by allowing for worker-employer match-specific productivity, whereby a specific type of worker may be more productive in some firms compared to others.

3.1. **A single firm case: Pure monopsony.**

Initially introduced by Robinson (1933), monopsony was used to describe a fairly specialized labor market in which a single firm hired labor in an isolated labor market (analogous to a monopolist in a product market). More recently, the term has been used to describe a labor market in which firms face upward-sloping labor supply curves.[20] To develop intuition for wage determination, we start by analyzing a single firm economy ($J = 1$) and we focus on the interior solution of the firm maximization problem for simplicity. This is a useful benchmark since there are no strategic interactions and thus, we can examine the forces that affect wages in the model in a clear and transparency manner.

Under Assumption 3, the single firm optimal wage equation (2.8) becomes:

$$w_{k1} = \psi_{k1}\frac{\mathcal{E}_{k1}(w_{k1})}{1+\mathcal{E}_{k1}(w_{k1})}, \quad \text{with } \psi_{k1} \equiv \lambda_1\theta_1\gamma_{k1} \tag{3.1}$$

where the elasticity $\mathcal{E}_{k1}(w_{k1})$ is written as a function of w_{k1} to emphasize that (3.1) is an implicit equation. In order to analyze the effects of the primitives on the equilibrium wage we need to solve this implicit equation which depends on the functional form of the elasticity, which in turn depends

[20]Please see Manning (2003) for a more detailed discussion. Recent applications of this idea include Staiger et al (2010), Dube et al. (2018), and Kline et al. (2018).

on the distribution of the idiosyncratic part of workers preferences. Recall that the elasticity is defined as $\mathcal{E}_{k1} = \frac{w_{k1}}{s_{k1}} \frac{\partial s_{k1}}{\partial w_{k1}}$ where $s_{k1}(w_{k1})$ is given by:

$$s_{k1}(w_{k1}) = \mathbb{P}\Big(\frac{w_{k0}}{u_{k1}} e^{\epsilon_{k_i 0} - \epsilon_{k_i 1}} \leq w_{k1}\Big) \equiv \mathbb{P}\big(w^R_{ik1} \leq w_{k1}\big) \tag{3.2}$$

where $w^R_{ik1} \equiv \frac{w_{k0}}{u_{k1}} e^{\epsilon_{k_i 0} - \epsilon_{k_i 1}}$. Note that w^R_{ik1} can be interpreted as individual i's *reservation wage*, since $w^R_{ik1} < w_{k1} \Leftrightarrow U_{i0} < U_{i1}$. Therefore, individual i chooses to work for the monopsonist iff $w_{k1} > w^R_{ik1}$ (e.g., the offered wage exceeds reservation wage). In this case, the labor supply curve facing the firm *is* the cumulative distribution function (CDF) of reservation wages. This in turn depends on the distribution of $\epsilon_{i0} - \epsilon_{i1}$ in the population. We assume that firms cannot observe $\epsilon_{i0} - \epsilon_{i1}$ for any individual; rather they can only observe the distribution and $\frac{w_{k0}}{u_{k1}}$. Thus, they cannot perfectly price discriminate and workers earn rents due to the fact that they have private information about their reservation wage. Assuming that $\epsilon_{i0} - \epsilon_{i1}$ follows a four parameter Beta distribution, we show in Appendix A.1 that the wage-posting model we consider here can rationalize the equilibrium wage predicted by the ex post wage-bargaining model, as initially discussed in Manning (2011). However, in general, the solution of the wage-posting model will not coincide with the wage-bargaining model, see Manning (2011, Section 3.1 & 3.2). The isomorphism with the wage-bargaining equilibrium wage depends on the distribution imposed on ϵ. Hereafter, we will mainly focus on the commonly used generalized extreme value distribution which is very tractable in a multinomial choice model (see Card et al. (2018) and Lamadon, Mogstad and Setzler (2018)).

Therefore, we introduce the following assumption:

Assumption 4 (Logit Economy)**.** *The stochastic terms of workers preferences, $(\epsilon_{ij})_{(k,j)\in(\mathcal{K}\times\mathcal{J})}$ are assumed to be i.i.d. random variables distributed according to the Extreme Value Type-I (Gumbel) distribution.*[21]

Under Assumption 4, we can show that the EWF & DMF in the pure monopsony case are:

$$w_{k1} = \frac{w_{k0}}{u_{k1}} \Big\{ \Big(1 + \frac{\psi_{k1} u_{k1}}{w_{k0}}\Big)^{1/2} - 1 \Big\}, \tag{3.3}$$

$$s_{k1} = 1 - \Big(1 + \frac{\psi_{k1} u_{k1}}{w_{k0}}\Big)^{-1/2}. \tag{3.4}$$

The EWF illustrates the different sources of wage inequality. First, wages may vary across workers due to differences in skill, γ_{k1}. A higher skilled worker in this economy increases the marginal product and therefore firms pay a higher wage to attract more skilled workers. The second source of wage inequality is the term, $\frac{w_{k0}}{u_{k1}}$. While this term does not lead to heterogeneity in reservation wages per se (for a given worker type) and thus does not affect the slope of the labor curve, it does affect workers' willingness-to-accept a job offer at the firm and thus impacts the wage offer. Intuitively, reservation wages, and hence wage offers, are greater when either a) the utility of working at the firm is high, b) the utility of being unemployed is low, or c) the non-unemployment wage is low.

[21]The Extreme Value Type-I distribution is a special case of the Generalized Extreme Value distribution with location 0, scale 1 and shape parameter 0.

These all capture the role of compensating differentials and outside options. Lastly, heterogeneity in reservation wages is implicitly captured by the distribution of $\epsilon_{i0} - \epsilon_{i1}$ in the population which under the Gumbel assumption has a log-logistic distribution. Interestingly, both the EWFs derived under the beta distribution, i.e. eqs (A.1) and this present one capture the essential feature of the compensating differential theory, in the sense that the equilibrium wage of a type k worker directly depends on type k workers non-pecuniary's preference. This feature will allow us to analyze how different preferences for firm amenities may contribute to wage inequality. In the next section, we will analyze how strategic competition between firms will affect the EWFs and DMFs. However, before doing so, we analyze how our framework can be used to provide a causal explanation of equilibrium wage gap between equally productive workers in the pure monopsony case.

3.1.1. *Wage discrimination in presence of Monopsony power.* Robinson's original application of monopsony was developed to explain wage differentials between equally productive women and men. In order to shed light on this, consider two type of workers who have different observable characteristics — for instance, male vs female, black vs white, etc— but have the same ability/skill level $\tilde{k}$. To fix the ideas, let's denote them by $k = (m, \tilde{k})$ and $k' = (f, \tilde{k})$ and assume that we have a linear production function. In this case, the wage gap between these two equally productive workers can be written as follows:

$$\frac{w_{k1}}{w_{k'1}} = \underbrace{\frac{\gamma_{(m,\tilde{k})1}}{\gamma_{(f,\tilde{k})1}}}_{BWG} \times \underbrace{\frac{\left(\frac{\mathcal{E}_{k1}}{1+\mathcal{E}_{k1}}\right)}{\left(\frac{\mathcal{E}_{k'1}}{1+\mathcal{E}_{k'1}}\right)}}_{RWG}. \tag{3.5}$$

Therefore, the wage gap between two equally productive workers are driven by two main components, namely the "Robinsonian" and the "Beckerian" wage discrimination gap, respectively BWG and RWG.

As detailed in the Appendix A.2, in the Becker's model, because of taste-based discrimination, a firm can evaluate the marginal contributions of two different type of workers differently even if they have the same level of skill, $\gamma_{(m,\tilde{k})j} \neq \gamma_{(f,\tilde{k})j}$. Thus, this could lead to the "Beckerian" wage discrimination gap. However, for Robinson, in a monopsony market, the wage gap can be rationalized by the fact that two groups of workers may have different labor supply elasticities, i.e. $\mathcal{E}_{kj} \neq \mathcal{E}_{k'j}$.

Following Robinson's arguments, some have justified the presence of the gender wage gap (even for equally productive workers) by the relatively lower elasticity of firm-specific labor supply for women; see for instance Ransom and Oaxaca (2004) and Ransom and Lambson (2011). Our EWF & DMF in the Logit economy will permit a more causal interpretation of the determinants of the equilibrium wage gap. More precisely, under Assumptions 3 and 4, the equilibrium wage gap is:

$$\frac{w_{k1}}{w_{k'1}} = \frac{w_{k0}u_{k'1}\left\{\left(1 + \frac{\psi_{k1}u_{k1}}{w_{k0}}\right)^{1/2} - 1\right\}}{w_{k'0}u_{k1}\left\{\left(1 + \frac{\psi_{k'1}u_{k'1}}{w_{k'0}}\right)^{1/2} - 1\right\}}, \tag{3.6}$$

This equation provides a clear understanding of the primitive parameters that can contribute causally to the gender/race wage gap. More precisely, differences in the non-employment wage, preferences for amenities, and the way that firms perceive workers productivity due to distaste, i.e. γ_{k1} all contribute to wage inequality. If we consider that firms do not engage in taste-based discrimination, i.e. $\gamma_{k1} = \gamma_{k'1}$, then the EWG is explained by the fact that women/black have higher preferences for firms amenities and/or have also lower non-employment wages which causes their reservation wages to be lower. In this scenario, the model predicts (eq (3.4)) that we should have a higher relative proportion of women/black hired at the equilibrium, i.e. $s_{k'1} \geq s_{k1}$. If this is not what we observe, according to this model, there must be taste-based discrimination as argued by Becker, i.e. $BWG < 1$. In general, papers that seek to rationalize the gender wage gap do not link the wage gap with the gender employment gap, i.e. $\frac{s_{k1}}{s_{k'1}}$ while those two are fundamentally linked. One advantage of using the DWFs here is to analyze those two interdependent features of the market simultaneously. In our present context, using both allows to test for the presence of a BWG.

3.2. **Oligopsony.** Now, we consider that we have a finite number of firms $J \geq 2$. Under Assumption 4 it can be shown — see MacFadden (1978)— that $G_{k\cdot}(v_{k\cdot}) = \ln\left\{e^{v_{k0}} + \sum_{j=1}^{J} e^{v_{kj}}\right\}$ and the labor supply elasticity simplifies to:

$$\begin{aligned} \mathcal{E}_{kj} &= \big(1 - s_{kj}\big), && (3.7) \\ s_{kj} &= \frac{u_{kj} w_{kj}}{w_{k0} + \sum_{j \in \mathcal{J}} u_{kj} w_{kj}}. && (3.8) \end{aligned}$$

This shows that the labor supply elasticity depends on the market share. Intuitively, if the firm is more desirable to workers, this will increase labor supply to the firm and so the share will be higher and this will also imply that these workers are more inelastic.

Remark 3 (Bajari and Benkard (2003))**.** *The generalized extreme value (GEV) distribution imposed on ϵ in Assumption 4 has some important implications that are worth-noting. (a)* ***Lack of Perfect Substitutes:*** *Each firm almost surely does not have a perfect substitute even as $J \to \infty$. Each worker would suffer utility losses that are almost surely bounded away from zero if her first choice of firm is removed from the choice set. (b)* ***Lack of Perfect Competition:*** *The elasticity is always bounded, i.e. $\mathcal{E}_{kj} < \infty$ even if $J \to \infty$, then we never obtain the EWF in a model with perfect competition. However, these two restrictions can be relaxed using alternative distribution for ϵ.*

The main difference that emerges from having multiple firms is that the elasticity related to type k workers $\mathcal{E}_{kj}$ no longer depends only on one endogenous quantity w_{kj}, but now depends on all potential wage offers that type k workers could receive from all other competing firms. Let $\Delta_{k,-j} = w_{k0} + \sum_{j' \neq j} w_{kj} u_{kj}$ be an index that measures the aggregate gross mean utility of all other options except firm j. Generalizing eq (3.3) to the multiple firm case under Assumptions 3 and 4 leads to the following equation:

$$w_{kj} = \frac{\Delta_{k,-j}}{u_{kj}} \left\{ \left(1 + \frac{\psi_{kj} u_{kj}}{\Delta_{k,-j}}\right)^{1/2} - 1 \right\}, \quad \forall k \in \mathcal{C}^j, j \in \mathcal{J}. \qquad (3.9)$$

Under the assumption that $u_{kj}= u_k$, one can see that the firm benchmarks it's wage against the wages of other firms in the market, which serves as a kind of ex-ante outside option. In the Logit Economy, each firm has the same weight and affects the outside option symmetrically (under the assumption that the non-pecuniary term is the same across firms). With heterogeneous non-pecuniary preferences, the firm will place relatively higher weight on the wages of firms that the worker prefers.

Of course, unlike the single firm case, the RHS of the equation depends on wages which are *endogenously* set by others firms through the index $\Delta_{k,-j}$. As in the monopsony case, we consider that w_{k0} is a predetermined outcome —determined outside the market. More formally, we can state the following assumption:

Assumption 5. *Assume that $w_{k0}, k \in \mathcal{K}$ are pre-determined outcomes which are not endogenously determined within the market under study, i.e. they are not the result of either firms or workers optimization problem.*

Under Assumption 5, finding the EWFs and DMFs that depends only on w_{k0} and the structural parameters $u_{kj}, \theta_j, \lambda_j$, and γ_{kj} requires solving a non-linear system of $K \times J$ equations in $K \times J$ unknowns, i.e eq (3.9). Now, we will proceed by showing that solving the $K \times J$ non-linear system of equations is equivalent to solve K distincts non-linear equations where each equation has only one unknown. This significantly reduces the dimensionality of the problem since the numbers of equations remains fixed even if the number of firms J become arbitrarily large. This simplification will be mainly possible by providing an alternative tractable characterization of the equilibrium model that make use of the quasi-supply.

In the one-to-one marriage matching literature, Choo and Siow (2006) proposed to use the quasi-supply function instead of the supply function to obtain a parsimonious and tractable marriage matching function. We use the same idea and consider the quasi-supply $\frac{s_{kj}}{s_{k0}}$ for $s_{k0} > 0$ and then from equation (3.8) we obtain:

$$s_{kj} = s_{k0} \frac{u_{kj}}{w_{k0}} w_{kj}, \ \forall k \in \mathcal{C}^j, j \in \mathcal{J}. \tag{3.10}$$

Considering the quasi-supply relative to the unemployed share here is not innocuous, in fact according our ARUM we always have $s_{k0} > 0$ for all $k \in \mathcal{K}$. However, the quasi-supply relative to s_{kj^*} may not be always defined since the firm j^* may not hire all type k at the optimum, i.e. $\overline{\mathcal{C}^j} \equiv \mathcal{K} \setminus \mathcal{C}^j \neq \{\emptyset\}$. Now, in the Logit economy and under Assumption 3, equation (2.10) simplifies to

$$w_{kj} = \psi_{kj} \frac{1 - s_{kj}}{2 - s_{kj}}, \quad \forall k \in \mathcal{C}^j, j \in \mathcal{J}. \tag{3.11}$$

where $\psi_{kj} \equiv \lambda_j \theta_j \gamma_{kj}$. Plugging the Logit quasi-supply equation equation (3.10) into equation (3.11) and rearranging, we obtain the following quadratic equation in w_{kj}:

$$\frac{u_{kj}}{w_{k0}} s_{k0} (w_{kj})^2 - \Big(2 + \frac{u_{kj}}{w_{k0}} s_{k0} \psi_{kj}\Big) w_{kj} + \psi_{kj} = 0. \tag{3.12}$$

It can be showed that this equation has two non-negative solutions. Then, generating two potential wage streams $(w_{kj}, \tilde{w}_{kj})_{(k,j)\in(\mathcal{K}\times\mathcal{J})}$ and distributions of workers across firms $(\mu_{kj}, \tilde{\mu}_{kj})_{(k,j)\in(\mathcal{K}\times\mathcal{J})}$ that solve the firms F.O.C and are compatible with the Logit quasi-supply equations i.e. eq (3.11, and 3.10). However, it can be verified that one of the pair $(w_{kj}, \mu_{kj})_{(k,j)\in(\mathcal{K}\times\mathcal{J})}$ is the solution to the cost minimization problem since $\sum_{k=1}^{K} w_{kj}\mu_{kj} < \sum_{k=1}^{K} \tilde{w}_{kj}\tilde{\mu}_{kj}$. Therefore, the EWF and DMF in our many-to-one employee-employer matching model are given by this pair:

$$w_{kj} = \frac{1}{2}\psi_{kj} + \frac{w_{k0}}{s_{k0}u_{kj}} - \sqrt{\frac{1}{4}\psi_{kj}^2 + \Big(\frac{w_{k0}}{s_{k0}u_{kj}}\Big)^2}, \tag{3.13}$$

$$s_{kj} = \frac{1}{2}\psi_{kj}\frac{u_{kj}}{w_{k0}}s_{k0} + 1 - \sqrt{\Big(\frac{1}{2}\psi_{kj}\frac{u_{kj}}{w_{k0}}s_{k0}\Big)^2 + 1}, \tag{3.14}$$

for all $k \in \mathcal{C}^j, j \in \mathcal{J}$, and $w_{kj} = s_{kj} = 0$ for all $k \in \overline{\mathcal{C}^j} \equiv \mathcal{K} \setminus \mathcal{C}^j, j \in \mathcal{J}$.

Notice that unlike to the monopsony case, the EWF & DMF, i.e. eq (3.13, 3.14) depend on the equilibrium object s_{k0}, which, for each type $k \in \mathcal{K}$ is the solution to the following non-linear equation:

$$s_{k0} + J_k + \sum_{j\in\mathcal{J}_k}\left\{\frac{1}{2}\psi_{kj}\frac{u_{kj}}{w_{k0}}s_{k0} - \sqrt{\Big(\frac{1}{2}\psi_{kj}\frac{u_{kj}}{w_{k0}}s_{k0}\Big)^2 + 1}\right\} = 1, \tag{3.15}$$

where $\mathcal{J}_k \equiv \{j \in \mathcal{J} : k \in \mathcal{C}^j\}$ and $J_k = |\mathcal{J}_k|$. This latter restriction is due to the population constraint eq (2.11).

The main advantage of characterizing the EWFs and DMFs in terms of the unemployed share s_{k0} as the only endogenous objects, is that it allow a more tractable characterization of the equilibrium which allows (a) to easily analyze the existence and uniqueness of the our equilibrium matching model and (b) to derive very informative closed form comparatives statistics. This approach of using the quasi-supply instead of the supply to derive the EWFs & DMFs is new to the labor literature as far as we are aware.

Let $H(\varphi_{k\cdot}, w_{k0}, s_{k0}) \equiv s_{k0} + J_k + \sum_{j\in\mathcal{J}_k}\left\{\frac{1}{2}\psi_{kj}\frac{u_{kj}}{w_{k0}}s_{k0} - \sqrt{\Big(\frac{1}{2}\psi_{kj}\frac{u_{kj}}{w_{k0}}s_{k0}\Big)^2 + 1}\right\} - 1$ where $\varphi_{k\cdot} = (\psi_{k\cdot}, u_{k\cdot})$ with $\psi_{k\cdot} \equiv (\psi_{k1}, ..., \psi_{kJ})$, $u_{k\cdot} \equiv (u_{k1}, ..., u_{kJ})$. We have the following result:

Lemma 1. *For all $k \in \mathcal{K}$ and $\varphi_{k\cdot} \in \big(\mathbb{R}_{>0}\big)^{2J}$ we have the following results:*

(i) $H(\varphi_{k\cdot}, w_{k0}, 0) = -1$, *and* $H(\varphi_{k\cdot}, w_{k0}, 1) > 0$.
(ii) $\frac{\partial H(\varphi_{k\cdot}, w_{k0}, s_{k0})}{\partial s_{k0}} = 1 + \sum_{j\in\mathcal{J}_k}\frac{1}{2}\psi_{kj}\frac{u_{kj}}{w_{k0}}\zeta_{kj} > 1$,

where $\zeta_{kj} = 1 - \frac{\frac{1}{2}\psi_{kj}\frac{u_{kj}}{w_{k0}}s_{k0}}{\left\{\Big(\frac{1}{2}\psi_{kj}\frac{u_{kj}}{w_{k0}}s_{k0}\Big)^2 + 1\right\}^{1/2}} \in (0,1)$. *Then, there exists a unique $s_{k0} \in (0,1)$ that solve the equation (3.15).*

Lemma 1, shows that given a fixed set of parameters $\varphi = (\varphi_{1\cdot}, ..., \varphi_{K\cdot})$ and w_{k0} there is a unique $s_{k0} \in (0,1)$ that solves equation (3.15). This leads to the following result:

Theorem 1 (Existence and Uniqueness of the equilibrium)**.** *Consider a Logit economy where workers have preferences in the ARUM form and firms have linear productions functions (in other words Assumptions 3 and 4 hold), and Assumption 5 holds, then an equilibrium outcome of our many-to-one employee-employer matching model with imperfect competition always exists and is unique. This equilibrium outcome is entirely characterized by the EWF & DMF, eqs (3.13, 3.14) and the population constraints eq (2.11).*

Notice that if, from equation (3.15), we were able to get a closed-form solution of the unemployed share s_{k0} as a function of the structural parameters φ and w_{k0}, we could substitute them into the EWFs and DMFs and obtain an equilibrium wage and matching function in terms of these parameters. However, even if an analytic solution does not exist, we can apply the implicit function theorem to (3.15) and make use of our EWFs & DMFs, eqs (3.13, 3.14), to derive a set of comparative statics.

Assumption 6. *An exogenous change of amenities does not affect workers and firms productivity and vice versa, more precisely* $\frac{\partial \psi_{kj}}{\partial u_{kj}} = 0$.

Theorem 2 (Comparative Statics)**.** *Under the conditions of Theorem 1, and Assumption 6 and given* $\varphi_{k\cdot} \in \left(\mathbb{R}_{>0}\right)^{2(J+1)}$, $w_{k0} \in (0, \infty)$ *for all* $k \in \mathcal{K}$, *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 comparative statics hold:*

(i) *Type-specific elasticity of unemployed share.*

$$\frac{\psi_{kj}}{s_{k0}} \frac{\partial s_{k0}}{\partial \psi_{kj}} = \frac{u_{kj}}{s_{k0}} \frac{\partial s_{k0}}{\partial u_{kj}} = -\frac{\frac{1}{2}\psi_{kj}\frac{u_{kj}}{w_{k0}}\zeta_{kj}}{1 + \sum_{j \in \mathcal{J}_k} \frac{1}{2}\psi_{kj}\frac{u_{kj}}{w_{k0}}\zeta_{kj}} \in (-1, 0), \quad j, \in \mathcal{J}_k \tag{3.16}$$

$$\frac{w_{k0}}{s_{k0}} \frac{\partial s_{k0}}{\partial w_{k0}} = \frac{\sum_{j \in \mathcal{J}_k} \frac{1}{2}\psi_{kj}\frac{u_{kj}}{w_{k0}}\zeta_{kj}}{1 + \sum_{j \in \mathcal{J}_k} \frac{1}{2}\psi_{kj}\frac{u_{kj}}{w_{k0}}\zeta_{kj}} \in (0, 1), \tag{3.17}$$

where $\zeta_{kj} = 1 - \frac{\frac{1}{2}\psi_{kj}\frac{u_{kj}}{w_{k0}}s_{k0}}{\left\{\left(\frac{1}{2}\psi_{kj}\frac{u_{kj}}{w_{k0}}s_{k0}\right)^2 + 1\right\}^{1/2}} \in (0, 1)$

(ii) *Aggregate Elasticity of unemployed shares.*

$$\sum_{j \in \mathcal{J}_k} \frac{\psi_{kj}}{s_{k0}} \frac{\partial s_{k0}}{\partial \psi_{kj}} = \sum_{j \in \mathcal{J}_k} \frac{u_{kj}}{s_{k0}} \frac{\partial s_{k0}}{\partial u_{kj}} = -\frac{w_{k0}}{s_{k0}} \frac{\partial s_{k0}}{\partial w_{k0}} \tag{3.18}$$

(iii) *Type-specific elasticity of workers' shares.*

$$\frac{\psi_{kj'}}{s_{kj}} \frac{\partial s_{kj}}{\partial \psi_{kj'}} = \frac{u_{kj'}}{s_{kj}} \frac{\partial s_{kj}}{\partial u_{kj'}} = \begin{cases} \frac{s_{k0}w_{k0}u_{kj}\psi_{kj}}{2s_{kj}} \frac{u_{kj'}}{s_{k0}} \frac{\partial s_{k0}}{\partial u_{kj'}}\zeta_{kj} < 0, & j' \neq j \in \mathcal{J}_k \\ \frac{s_{k0}w_{k0}u_{kj}\psi_{kj}}{2s_{kj}} \left(1 + \frac{u_{kj}}{s_{k0}} \frac{\partial s_{k0}}{\partial u_{kj}}\right)\zeta_{kj} > 0, & j' = j \in \mathcal{J}_k \end{cases} \tag{3.19}$$

$$\frac{w_{k0}}{s_{kj}} \frac{\partial s_{kj}}{\partial w_{k0}} = -\frac{s_{k0}w_{k0}u_{kj}\psi_{kj}}{2s_{kj}} \left(1 - \frac{w_{k0}}{s_{k0}} \frac{\partial s_{k0}}{\partial w_{k0}}\right)\zeta_{kj} < 0, \tag{3.20}$$

(iv) *Type-specific elasticity of workers' wages.*

$$\frac{\psi_{kj'}}{w_{kj}}\frac{\partial w_{kj}}{\partial \psi_{kj'}} = \frac{u_{kj'}}{w_{kj}}\frac{\partial w_{kj}}{\partial u_{kj'}} = -\frac{w_{k0}}{s_{k0}w_{kj}u_{kj}}\frac{\psi_{kj'}}{s_{k0}}\frac{\partial s_{k0}}{\partial \psi_{kj'}}\eta_{kj} > 0, \quad j' \neq j \in \mathcal{J}_k, \tag{3.21}$$

$$\frac{u_{kj}}{w_{kj}}\frac{\partial w_{kj}}{\partial u_{kj}} = -\frac{w_{k0}}{s_{k0}w_{kj}u_{kj}}\Big(1 + \frac{u_{kj}}{s_{k0}}\frac{\partial s_{k0}}{\partial u_{kj}}\Big)\eta_{kj} < 0, \tag{3.22}$$

$$\frac{\psi_{kj}}{w_{kj}}\frac{\partial w_{kj}}{\partial \psi_{kj}} = -\frac{w_{k0}}{s_{k0}w_{kj}u_{kj}}\Big(\frac{u_{kj}}{s_{k0}}\frac{\partial s_{k0}}{\partial u_{kj}}\Big)\eta_{kj} + \frac{\psi_{kj}}{2w_{kj}}\varkappa_{kj} > 0, \tag{3.23}$$

$$\frac{w_{k0}}{w_{kj}}\frac{\partial w_{kj}}{\partial w_{k0}} = \frac{w_{k0}}{s_{k0}w_{kj}u_{kj}}\Big(1 - \frac{w_{k0}}{s_{k0}}\frac{\partial s_{k0}}{\partial w_{k0}}\Big)\eta_{kj} > 0 \tag{3.24}$$

$$\textit{where } \eta_{kj} = 1 - \frac{\frac{w_{k0}}{s_{k0}u_{kj}}}{\left\{\frac{1}{4}\psi_{kj}^2 + \left(\frac{w_{k0}}{s_{k0}u_{kj}}\right)^2\right\}^{1/2}} \in (0,1) \textit{ and } \varkappa_{kj} = 1 - \frac{\frac{1}{2}\psi_{kj}}{\left\{\frac{1}{4}\psi_{kj}^2 + \left(\frac{w_{k0}}{s_{k0}u_{kj}}\right)^2\right\}^{1/2}} \in (0,1).$$

(iii) *Variations in type-specific labor supply curves.*

$$\frac{\partial \mathcal{E}_{kj}}{\partial u_{kj'}} = \begin{cases} -\frac{\partial s_{kj}}{\partial u_{kj'}} > 0, & j' \neq j \in \mathcal{J}_k \\ -\frac{\partial s_{kj}}{\partial u_{kj}} < 0, & j' = j \in \mathcal{J}_k \end{cases}$$

$$\frac{\partial \mathcal{E}_{kj}}{\partial \psi_{kj'}} = \begin{cases} -\frac{\partial s_{kj}}{\partial \psi_{kj'}} > 0, & j' \neq j \in \mathcal{J}_k \\ -\frac{\partial s_{kj}}{\partial \psi_{kj}} < 0, & j' = j \in \mathcal{J}_k \end{cases}$$

$$\frac{\partial \mathcal{E}_{kj}}{\partial w_{k0}} = -\frac{\partial s_{kj}}{\partial w_{k0}} > 0 \tag{3.25}$$

It is noteworthy that many of the comparative statics carry over from the pure monopsony case. First, wages at the firm are increasing with respect to worker productivity at the firm, $\frac{\partial w_{kj}}{\partial \psi_{kj}} > 0$. Second, wages are increasing with respect to the non-employment wage $\frac{\partial w_{kj}}{\partial w_{k0}} > 0$ and decreasing with respect to the amenity $\frac{\partial w_{kj}}{\partial u_{kj}} < 0$. A similar set of results obtains for the market shares. Consistent with the pure monopsony case, we also see that there is a symmetry in how u_{kj} and ψ_{kj} affects s_{kj}, although this symmetry does not carry over to the EWF, again similar to the pure monopsony case. Thus, the presence of strategic interactions does not modify the sign of these comparative statics.

In terms of how the productivities and amenities at *competing* firms impact the wage and share at a given firm, note that all of these effects operate through s_{k0}, the unemployment share. This is decreasing in both firm productivity and firm amenities as can be seen in equation (3.16). Thus, we see that an increase in *either* productivity or the amenity of firm j' increases the share at that firm and reduces it at firm j (see equation (3.19)). There is also a symmetry in how the productivity and amenities of other firms affects the market share at a given firm.

In terms of wages at a firm, we see they are increasing in productivity at other firms $\frac{\partial w_{kj}}{\partial \psi_{kj'}} > 0$. Intuitively, this increases wages at the competing firm and thus increases the labor supply elasticity to the firm which leads to a higher wage, as can be seen above. This implies that if there is a positive idiosyncratic shock to the productivity of a firm, wages are "strategic complements" – e.g., if firm j' increases its wage for a given worker type since productivity has increased, the best response is for firm j to increase the wage. On the other hand, wages at the firm are decreasing in the amenities of other firms $\frac{\partial w_{kj}}{\partial u_{kj'}} > 0$. Thus, if there is an exogenous increase in amenities at firm j', wages are "strategic substitutes" – wages at firm j' decrease but wages at firm j increase. The intuition for this is that the increase in amenities at the competing firm is greater than the reduction in the wage at that firm on net and this acts to increase the labor supply elasticity to the firm leading to a higher wage.

Thus, whether wages are strategic complements depends on what causes wages to change in the first place. Nonetheless, these comparative statics results are quite different from those that would obtain in a model of monopsonistic competition where wage setting is not affected by the firm's competitors.

Remark 4. *It is worth noting that Assumption 6 is mainly assumed for simplicity. In fact, if the sign of* $\frac{\partial \psi_{kj}}{\partial u_{kj}}$ *is known ex-ante, for example if we are in a context where an exogenous increase of amenities is believed to improve workers productivity, that is* $\frac{\partial \psi_{kj}}{\partial u_{kj}} > 0$*, we can re-do the comparative statistics under this condition. For instance, in such a case an exogenous increase of amenities will not necessarily leads to a decrease of wage as predicted in eq (3.23).*

3.3. **Social Surplus, Generalized Entropy and Market Concentration.** We define the *two-sided social surplus* function of our many-to-one matching model with imperfect competition as follows:

$$\mathcal{W} = \sum_{k=1}^{K} m_k G_{k\cdot}(v_{k\cdot}) + \sum_{j=1}^{J} \Big(F^j(\mu_{\cdot j}) - \sum_{k=1}^{K} w_{kj}\mu_{kj} \Big) \tag{3.26}$$

where the first term, namely the social surplus function $\mathcal{G} \equiv \sum_{k=1}^{K} m_k G_{k\cdot}(s_{k\cdot})$ is the summation of the expected utilities of all types of workers in the market —in other terms the aggregate of all type k social surplus function— and the second term represents the summation of the profit of all the firms in the market. Under the linear production function assumption, i.e. Assumption 3, the surplus function simplifies to:

$$\mathcal{W} = \sum_{k=1}^{K} \Big[m_k G_{k\cdot}(v_{k\cdot}) + \sum_{j=1}^{J} (\gamma_{kj} - w_{kj}) s_{kj} \Big]. \tag{3.27}$$

Let $G^*_{k\cdot}(s_{k\cdot})$ denotes the *convex conjugate* or *Legendre-Fenchel transform* of $G_{k\cdot}(s_{k\cdot})$. The convex duality implies the following relationship between the social surplus function and its convex conjugate[22]:

$$G_{k\cdot}(s_{k\cdot}) = \sum_{j=0}^{J} v_{kj} s_{kj} - G^*_{k\cdot}(s_{k\cdot}). \tag{3.28}$$

Using the above relationship (3.28), the two-sided social surplus function becomes:

$$\mathcal{W} = \sum_{k=1}^{K} m_k \Big[\sum_{j=0}^{J} (v_{kj} + \gamma_{kj} - w_{kj}) s_{kj} \Big] - \sum_{k=1}^{K} m_k G^*_{k\cdot}(s_{k\cdot}). \tag{3.29}$$

[22]Please see Galichon and Salanié (2015) for more detailed discussion

This last formula shows that the two-sided social surplus function is the summation of two main components: (i) a summation of the deterministic gains obtained in the equilibrium matching by all types k workers and all the firms and (ii) a measure of randomness existing in the market. This second term is essentially due to the unobserved heterogeneity on the workers utilities. When ϵ follows the logit distribution (Assumption 4), $-G^*_{k\cdot}(s_{k\cdot})$ is the usual entropy, which in information theory is considered as a natural measure of statistical disorder. Following Galichon and Salanié (2015) we will denote $-G^* \equiv \sum_{k=1}^K m_k G^*_{k\cdot}(s_{k\cdot})$ the generalized entropy. Notice that G^* captures the level of *incomplete information* we have in the market. We will show that the entropy can be expressed as a weighted function of the Hannah-Kay (1971) type k workers employment measure of market concentration. In other terms the entropy is the natural measure of concentration in our model. First of all, notice that workers are differentiate by types and we could have more concentration for some type of workers than others. So, we can consider to define an index concentration by worker types. The Hannah-Kay (1971) general index measure of concentration of type k worker is defined as:

$$HK_\alpha(s_{k\cdot}) = \begin{cases} \left(\sum_j^J s_{kj}^\alpha\right)^{\frac{1}{\alpha-1}} & \text{if } \alpha > 0, \alpha \neq 1,^{23} \\ \Pi_{j=0}^J s_{kj}^{s_{kj}} & \text{if } \alpha = 1. \end{cases} \tag{3.30}$$

It can be shown that under Assumption 4,

$$G^*_{k\cdot}(s_{k\cdot}) = \sum_{j=0}^J s_{kj} \ln s_{kj} = \ln HK_1(s_{k\cdot}). \tag{3.31}$$

As can be seen in our model, both the social welfare and the social surplus functions directly depend on the Hannah-Kay (1971) measure of concentration:

$$\mathcal{W} = \sum_{k=1}^K m_k \Big[\sum_{j=0}^J (v_{kj} + \gamma_{kj} - w_{kj}) s_{kj}\Big] - \sum_{k=1}^K m_k \ln HK_1(s_{k\cdot}), \tag{3.32}$$

$$\mathcal{G} = \sum_{j=0}^J v_{kj} s_{kj} - \sum_{k=1}^K m_k \ln HK_1(s_{k\cdot}). \tag{3.33}$$

Notice that $HK_1(s_{k\cdot})$ reaches it lower bound $1/(J+1)$ when $s_{kj} = 1/J+1$, meaning that type k workers are uniformly distributed over the $J+1$ firms considering that unemployment is firm 0, while it reaches its upper bound 1 when only one firm hires all type k workers. When the $HK_1(s_{k\cdot}) = \frac{1}{J+1}$ for all k, we mimic the perfect competition while the second ones is achieved in a pure monopsony where all workers are employed. Remark that an uniform increase of the $HK_1(s_{k\cdot})$ employment concentration measure decreases the generalized entropy part of both surplus functions but does not lead necessarily to a decrease of the surplus function because of the workers heterogeneity in both they skills γ_{kj} and non-pecuniary preferences u_{kj}.

Moreover, notice that policy changes or exogenous shocks may affect non uniformly the concentration index, in the sense that the market could be more concentrated for some workers rather than others under policy change. In the following our estimation of u_{kj} and the fact that the equilibrium is unique will allow us to simulate the impact of different changes on the concentration index and both social surplus functions.

4. Rationalizing AKM's wage equation and its extensions

In this section, we derive the conditions that are necessary and sufficient to rationalize the AKM wage equation and some of its recent extensions in an oligopsony market. The empirical literature analyzing employer-employee wage data has been widely influenced by AKM who decompose log earnings into components associated with *time invariant* unobserved worker and firm heterogeneity. The AKM decomposition has been used to analyze firms and workers contribution to overall earnings inequality. This section considers whether the AKM reduced-form wage equation can be rationalized by our structural model. The wage equation (3.11) in the Logit economy with perfectly substitutable workers can be written as

$$\log w_{kj} \quad = \quad \log \delta_j + \log \gamma_{kj} + \log \frac{1 - s_{kj}}{2 - s_{kj}}$$

where $\delta_j \equiv \lambda_j \theta_j$. Thus, a sufficient condition for the Logit Economy to deliver the AKM wage structure is $\gamma_{kj} = \gamma_k$ (no complementarities between workers and firms), and constant elasticities. In models of monopsonistic competition considered in the literature, the elasticity is constant. However, in our oligopsony model with the Logit economy or a Nested Logit economy —see eqs (3.8, 3.7) for the elasticity in the Logit economy and eq (A.6) in Appendix A.3 for the Nested Logit Economy— the elasticities are variable since they generally depend on the wages set by other firms. In order to examine how AKM can arise within our structural model, we consider an alternative distribution of the error terms such that the following holds.

Assumption 7. *Assume that the elasticities of the workers' labor supply have the following functional form:* $\mathcal{E}_{kj}(w) = \frac{w_{kj}}{w_{kj} - \pi_{kj}} \beta_{kj}$ *such that* $w_{kj} > \pi_{kj}$ *where* π_{kj} *is not a function of* w_{kj} *and* $\beta_{kj} > 0$ *is a positive constant, and this for all* $(k, j) \in (\mathcal{K} \times \mathcal{J})$.

Solving the firms' optimal wage equations (2.8) under Assumptions 3 and 7 provides the following EWFs:

$$w_{kj} = \frac{\beta_{kj}}{1 + \beta_{kj}} (\lambda_j \theta_j \gamma_{kj}) + \frac{1}{1 + \beta_{kj}} \pi_{kj}, \quad j \in \mathcal{J}_k. \tag{4.1}$$

If $\pi_{kj} = \frac{w_{k0} u_{k0}}{u_{kj}} a$ one can see that this generalizes the "wage-bargaining" form of the EWF derived earlier, i.e. eq (A.1) to the case of multiple firms.

Relation with AKM, and Bonhomme, Lamadon and Manresa (2017). If (i) $\pi_{kj} = 0$ either because type k workers have no unemployment benefits ($w_{k0} \to 0$) or have a very high disutility of being unemployed ($u_{k0} \to 0$ then $\ln u_{k0} \to -\infty$), and (ii) β_{kj} the resulting labor supply elasticity is only firm-specific, i.e. $\beta_{kj} = \beta_j$ the EWF becomes:

$$\ln w_{kj} = \ln \delta_j + \ln \gamma_{kj} \tag{4.2}$$

where $\delta_j \equiv \lambda_j \theta_j \frac{\beta_j}{1+\beta_j}$. AKM and the wage equation in Bonhomme, Lamadon and Manresa (2017) can be seen as special cases. If one uses the following multiplicative specification $\ln \gamma_{kj} = \Delta_k \xi_j$, when $\xi_j = 1$ we recover the log additive form of the AKM model and when $\xi_j \neq 1$ we recover the multiplicative complementarity used in Bonhomme et al (2017). The upshot is that its possible to obtain a constant elasticity in our structural model for a suitable distribution of the error terms.

Assumption 7 is a "high level" assumption on the labor supply elasticity and a key question is whether it could be rationalized by some random utility maximization problem. In other words, does there exist a distribution of ϵ such that workers stochastic utility maximization could rationalize the elasticities in Assumption 7? Before answering this question, let's introduce the following definition:

Definition 2. *[Local compatibility with the ARUM] For any fixed type $k \in \mathcal{K}$, workers set of choice probabilities $s_{kj}(v_{k\cdot})$, $j \in \mathcal{J})$ is locally compatible with stochastic utility maximization on a set $\mathcal{V}_{k\cdot} \subseteq \mathbb{R}^{J+1}$ if for $(r, j) \in (\mathcal{R} \times \mathcal{J})$, and all $v_{k\cdot} \in \mathcal{V}_{k\cdot}$ we can write*

$$s_{kj}(v_{k\cdot}) = \mathbb{P}\Big(\max_{\{(l \neq j) \in \mathcal{J}\}} \{v_{kl} + \epsilon_{kl}\} < \epsilon_{kj} + v_{kj} \Big) \tag{4.3}$$

where $\epsilon_{k\cdot} = (\epsilon_{k0}, \epsilon_{k1}, ..., \epsilon_{kJ})'$ is a stochastic $J+1$ vector that respect Assumption 1.

Whenever $\mathcal{V}_{k\cdot} = \mathbb{R}^{J+1}$ we have global compatibility as discussed in William (1977), Daly and Zachary (1979). Definition 2 is a specialization to our context of the local compatibility with an ARUM stochastic maximization due to Borsch-Supan (1990) and refined in Koning and Ridder (1994, 2003). Recall that $v_{kj} \equiv \ln w_{kj} u_{kj}$, $v_{k0} \equiv \ln w_{k0}$ which is the deterministic part of workers' utilities. Let us denote $\widetilde{\mathcal{V}}_{k\cdot}(a, b) = \Big\{ v_{k\cdot} : \ln a \leq v_{kj} - v_{k0} < \ln b, \quad j \in \mathcal{J} \Big\} = \Big\{ w_{k\cdot} : a \leq \frac{w_{kj} u_{kj}}{w_{k0} v_{k0}} < b, \quad j \in \mathcal{J} \Big\}$, where $0 \leq a < b < \infty$. We have the following result:

Proposition 2. *For a pair of finite constant (a, b) such that $0 \leq a < b < \infty$, a given type $k \in \mathcal{K}$, and $\beta_{kj} > 0$ the following set of workers choice probabilities*

$$s_{kj}(v_{k\cdot}) = \chi_a(k) \Big(e^{v_{kj} - v_{k0}} - a \Big)^{\beta_{kj}} = \chi_a(k) \Big(\frac{w_{kj} u_{kj}}{w_{k0} u_{k0}} - a \Big)^{\beta_{kj}}, \quad j \in \mathcal{J}_k, \tag{4.4}$$

$$s_{kj}(v_{k\cdot}) = 0, \quad j \notin \mathcal{J}_k, \tag{4.5}$$

$$s_{k0}(v_{k\cdot}) = 1 - \sum_{j \in \mathcal{J}_k} s_{kj}(v_{k\cdot}), \tag{4.6}$$

where $\chi_a(k) = \Big\{ \sum_{j \in \mathcal{J}_k} (b - a)^{\beta_{kj}} \Big\}^{-1}$ is locally compatible with stochastic utility maximization on the set $\widetilde{\mathcal{V}}_{k\cdot}(a, b)$. Moreover, the type k workers labor supply elasticity related to this set of choices is:

$$\mathcal{E}_{kj}(w) = \frac{w_{kj}}{w_{kj} - \frac{w_{k0} u_{k0}}{u_{kj}} a} \beta_{kj}, \quad j \in \mathcal{J}_k. \tag{4.7}$$

The proof of Proposition 2 is relegated in Appendix D.1. It is important to note that the constructed distribution does not respect the global compatibility with the ARUM. As pointed in Borsch-Supan (1990) and Koning and Ridder (1994, 2003) the global compatibility is very strong and may not be so important in practice. Indeed, in general, wages are observed only on a bounded support, so one cannot test the global compatibility with the ARUM outside the observed support.

Notice that Proposition 2 does not really restrict the support of the wage but imposes a bounded support of the ratio between the mean gross utilities in working in firm j vs being unemployed, i.e. $0 < a \leq \frac{w_{kj}u_{kj}}{w_{k0}v_{k0}} < b < \infty$.

Corollary 1. *The EWF eq (4.1) that generalized the wage-bargaining and AKM wage equation can be rationalized by a many-to-one structural matching model where a continuum of heterogenous workers maximize their stochastic preferences — which are in the ARUM form— and a finite number of heterogenous firms set wages by minimizing production costs in the presence of linear productions functions.*

4.1. **Equilibrium wages when workers are not gross-substitutes.** So far, we have considered the case where workers are perfect substitutes. As pointed out in the introduction a large part of the literature of many-to-one matching assumes that workers are gross-substitutes. A rare exception is Pycia (2012) but this paper does not consider unobserved heterogeneity and incomplete information as we entertain here. Now, we want to investigate cases where the firm's production function exhibits increasing return to scale which rules out the gross substitutes assumption. Please refer to Kelso and Crawford (1982, section 6) for a detailed discussion on the gross-substitute condition. For the sake of simplicity we focus on the case the labor supplies elasticities are exogenous. As a special case of Proposition 2, we consider that workers maximize a vector of stochastic utility of the ARUM form which generates the following set of choice probabilities:

$$s_{kj}(v_{k\cdot}) = \chi_0(k)\Big(\frac{w_{kj}u_{kj}}{w_{k0}u_{k0}}\Big)^{\beta_j} 1\{j \in J_k\}, \tag{4.8}$$

$$s_{k0}(v_{k\cdot}) = 1 - \sum_{j\in\mathcal{J}_k} s_{kj}(v_{k\cdot}). \tag{4.9}$$

Let's consider a more general form of production function:

Assumption 8. *The production function of firm j is given by:*

$$F^j(\mu_{\cdot j}) = \theta_j\Big(\sum_{k=1}^{K}\gamma_{kj}\mu_{kj}\Big)^{\alpha_j} = \theta_j\Big(L_j\Big)^{\alpha_j}, \tag{4.10}$$

where L_j denotes firm j's total labor efficiency, $\alpha_j = \frac{L_j}{F^j(\mu_{\cdot j})}\frac{\partial F^j(\mu_{\cdot j})}{\partial L_j} > 0$ is firm j's output elasticity respect to L_j.

As can be seen, as $\alpha_j > 1$(resp. $\alpha_j < 1$), firm'j production function exhibits increasing return to scale (res. decreasing return to scale) and constant return to scale when $\alpha_j = 1$. After derivations relegated to Appendix A.4, the following EWF & DMF:

$$w_{kj} = \Big(\lambda_j\theta_j\frac{\alpha_j\beta_j}{1+\beta_j}\Big)^{\frac{1}{1+(1-\alpha_j)\beta_j}}\gamma_{kj}\left\{\sum_{k=1}^{K} m_k\chi_0(k)(\gamma_{kj})^{1+\beta_j}\Big(\frac{u_{kj}}{w_{k0}}\Big)^{\beta_j}\right\}^{-\frac{(1-\alpha_j)}{1+(1-\alpha_j)\beta_j}} \tag{4.11}$$

$$s_{kj} = \Big(\lambda_j\theta_j\frac{\alpha_j\beta_j}{1+\beta_j}\Big)^{\frac{\beta_j}{1+(1-\alpha_j)\beta_j}}\chi_0(k)\Big(\frac{\gamma_{kj}u_{kj}}{w_{k0}}\Big)^{\beta_j}\left\{\sum_{k=1}^{K} m_k\chi_0(k)(\gamma_{kj})^{1+\beta_j}\Big(\frac{u_{kj}}{w_{k0}}\Big)^{\beta_j}\right\}^{-\frac{(1-\alpha_j)\beta_j}{1+(1-\alpha_j)\beta_j}} \tag{4.12}$$

for all $j \in \mathcal{J}_k$, and $w_{kj} = s_{kj} = 0$ for all $j \notin \mathcal{J}_k$. Thus, if we consider that workers have preferences in the ARUM form that rationalize the choice probabilities (4.8 & 4.9) and assumption 8 holds then, the EWFs and DMFs (4.11, 4.12) entirely characterize the unique equilibrium of the many-to-one matching model. Notice that the EWFs and DMFs are uniquely determined by only the primitives of the model. Therefore, they can be directly used to perform comparatives statics, since they depend only on exogenous parameters.

5. Econometric model: Identification and Estimation

In this section we focus on identification of the following structural parameters: u_{kj} and $\psi_{kj} \equiv \lambda_j \theta_j \gamma_{kj}$ in the Logit Economy. The term u_{kj} captures the non-pecuniary preference that type k workers have for firm j relative to the type k non-pecuniary preference of being unemployed. ψ_{kj} is a product of three terms that cannot be identified separately without further restrictions. Recall λ_j is the marginal cost of output. At the profit maximizing level of output, λ_j is equal to marginal revenue product, which leads to various scenarios. For instance, when firm j is a price taker in its output market, we have $\lambda_j = P_j$ where P_j is the exogenous price; however, in the presence of a downward-sloping firm-specific product demand curve, we have for each $\epsilon > 1$, $\lambda_j(\epsilon) = (\epsilon - \frac{1}{\epsilon})P_j(Y^j)^{-\frac{1}{\epsilon}}$. Therefore, if the researcher has more information on firm j marginal revenue product and also on P_j, the identification of ψ_{kj} may allow one to recover $\theta_j \gamma_{kj}$. Also, the of role match-specific productivity has been a main concern in this literature. The standard AKM wage equation rules out complementarities between workers and firms. The identification of ψ_{kj} will therefore be informative about the existence of such complementarities.

5.1. **When firms set wages based only on workers' observable types.** In this subsection, we consider that the main characteristics used by firms to set wages are also observed by the econometrician. A simple way to formalize this assumption in our model is the following:

Assumption 9. *Assume that*

$$u_{kj} = u_{\bar{k}j} \text{ and } w_{kj} = w_{\bar{k}j} \quad \forall \ (k, j) \in (\mathcal{K} \times \mathcal{J} \cup \{0\}). \tag{5.1}$$

This assumption has been used for identification and testability in various two-sided matching models. For instance, it has been considered in the seminal contribution of Choo and Siow (2006) and its main extensions such as Galichon and Salanié (2015), Choo (2015), Chiappori, Salanié, and Weiss (2017), Mourifié and Siow (2017), and Galichon, Kominers and Weber (2018) amongst others. These papers consider that the non-stochastic part of individuals' preferences depends only on the demographics of partners that are also observable to the econometrician. See also Fox (2018, Assumption 1), Graham (2011, section 3.1.3), Echenique et al (2013) and related discussions therein. A similar restriction has also been used in recent labor matching models; see for instance, Dupuy and Galichon (2018), Dupuy et al (2018), and Caldwell and Danielli (2018). Assumption 9 can be viewed as the two-sided version of the "matching/selection on observables" assumption widely used in the causal inference literature. While quite restrictive, this assumption has been key in most of the empirical papers trying to estimate structural parameters in two-sided matching models. In the following, we analyze the identifying power of this assumption in our framework and we later provide an alternative identification approach when this assumption fails to hold.

Proposition 3 (Identification)**.** *Whenever Assumptions 3, 4, and 9 hold we have the following identification result:*

$$u_{\bar{k}j} = \frac{w_{\bar{k}0}}{w_{\bar{k}j}} \frac{s_{\bar{k}j}}{(1 - \sum_{j \in \mathcal{J}_{\bar{k}}} s_{\bar{k}j})}, \quad \bar{k} \in \mathcal{C}^j, j \in \mathcal{J} \tag{5.2}$$

$$\psi_{\bar{k}j} = w_{\bar{k}j} \frac{2 - s_{\bar{k}j}}{1 - s_{\bar{k}j}}, \quad \bar{k} \in \mathcal{C}^j, j \in \mathcal{J} \tag{5.3}$$

where $w_{\bar{k}j}(s_{\bar{k}j})$ are the observed aggregate equilibrium wage (share) of type $\bar{k}$ individuals working in the firm j, and $w_{\bar{k}0}$ is the aggregate unemployment benefit of type $\bar{k}$ individuals.

Proposition 3 says that in a Logit economy where workers are perfect substitutes and firms set wages based on characteristics that are commonly observed by the econometrician, workers firm-specific (relative) non-pecuniary preferences for firms amenities and productivities (up to a multiplicative constant) are point identified using data from a single labor market. Notice that the identification results are obtained using the first order conditions derived earlier, i.e. (3.10, 3.11). Notice that we can recover $u_{\bar{k}j}$ and $\psi_{\bar{k}j}$ only for types $\bar{k}$ that have been hired by firm j, i.e. $\bar{k} \in \mathcal{C}^j$. Remark that $\bar{k} \notin \mathcal{C}^j$ can be rationalized by at least two different situations: (a) the firm's optimal wage for those types is $w_{\bar{k}j} = 0$ (e.g., a corner solution); in such a case, according our ARUM specification, no one would like to work at this wage for firm j, and this holds for all values of $u_{\bar{k}j} \in [0, \infty)$; (b) $u_{\bar{k}j} \to 0$, then under Assumption 1 (ii) this type $\bar{k}$ specific labor supply function for firm j is identically equal to zero and thus holds irrespective of the wage that the firm j would like to offer to them, $w_{\bar{k}j} \in [0, \infty)$. Both of these situations with very different values of non-pecuniary preferences would equivalently explain that we observe at the equilibrium, zero shares between certain types of workers and firms. Since the econometrician has no additional information that allows one to distinguish between those two scenarios, $u_{\bar{k}j}$ for all $\bar{k} \notin \mathcal{C}^j$ are fundamentally non-identified in the model we entertain here.

Similarly to Graham (2011), we introduce the following measure of complementarity, i.e. *Aggregate Local Complementarity* (ALC)

$$\phi_{\bar{k},j,\bar{k}',j'} = \ln \frac{\gamma_{\bar{k}j}\gamma_{\bar{k}'j'}}{\gamma_{\bar{k}j'}\gamma_{\bar{k}'j}}; \quad \bar{k}, \bar{k}' \in \mathcal{C}^j \cap \mathcal{C}^{j'} \tag{5.4}$$

where $\phi_{\bar{k},j,\bar{k}',j'}$ measures the aggregate difference between the incremental productivity return associated with hiring a type $\bar{k}$ versus $\bar{k}'$ worker across firm j versus j'. $\phi_{\bar{k},j,\bar{k}',j'}$ is a local measure of complementarity. When there is an ordering on the workers and firms types, $\phi_{\bar{k},j,\bar{k}',j'}$ is informative about the presence of (or lack theorof) assortative matching. Proposition 3 leads to the following corollary:

Corollary 2 (Identification of ALC)**.** *Whenever Assumptions 3, 4, and 9 hold the ALC is point identified*

$$\phi_{\bar{k},j,\bar{k}',j'} = \ln \frac{w_{\bar{k}j}w_{\bar{k}'j'}}{w_{\bar{k}j'}w_{\bar{k}'j}} + \ln\Big(\frac{2-s_{\bar{k}j}}{1-s_{\bar{k}j}}\Big)\Big(\frac{2-s_{\bar{k}'j'}}{1-s_{\bar{k}'j'}}\Big)\Big(\frac{1-s_{\bar{k}'j}}{2-s_{\bar{k}'j}}\Big)\Big(\frac{1-s_{\bar{k}j'}}{2-s_{\bar{k}j'}}\Big),\ \bar{k},\bar{k}' \in \mathcal{C}^j \cap \mathcal{C}^{j'} \quad (5.5)$$

As can be seen, even if $\gamma_{\bar{k}j}$ is not point identified, we can point identify the measure of complementarity related to it. Notice that if $k > k'$ and $j > j'$, $\phi_{\bar{k},j,\bar{k}',j'} > 0$ (resp. $\phi_{\bar{k},j,\bar{k}',j'} < 0$) implies positive (resp. negative) assortative matching, in the sense that highest type firms are assigned to the highest type workers. However, while we may have a clear ordering on workers types —education, age, etc.— it is unclear which criteria we should use to order firms, see Sorkin (2017). However, notice that under Assumption 9 the identification of the non-pecuniary part of workers' utilities $u_{\bar{k}j}$ provides us with a way to rank firms according worker preferences for firms amenities. We can therefore define

$$u_j^{np} \equiv \frac{1}{|\mathcal{C}^j|}\sum_{\bar{k}\in\mathcal{C}^j} u_{\bar{k}j} = \frac{1}{|\mathcal{C}^j|}\sum_{\bar{k}\in\mathcal{C}^j} \frac{w_{\bar{k}0}}{w_{\bar{k}j}} \frac{s_{\bar{k}j}}{(1-\sum\limits_{j\in\mathcal{J}_{\bar{k}}} s_{\bar{k}j})}, \quad (5.6)$$

where u_j^{np} captures the normalized aggregate preferences that all workers have for firm j amenities. The term $\frac{1}{|\mathcal{C}^j|}$ is used to suppress the impact of the firms' sizes on the ranking.

5.1.1. *Counterfactual analysis: An Inverse Identification Problem.* As Proposition 3 shows, for an observed distribution of aggregate equilibrium wages, matches $(w_{\bar{k}j}, s_{\bar{k}j})_{(\bar{k},j)\in(\mathcal{K}\times\mathcal{J})}$, and unemployment benefits $(w_{\bar{k}0})_{\bar{k}\in\mathcal{K}}$, we obtain a unique vector of non-pecuniary workers preferences and firm-specific productivities, i.e. $(u_{\bar{k}j}, \psi_{\bar{k}j})_{\bar{k}\in\mathcal{C}^j, j\in\mathcal{J}}$. In other words, based on aggregate observables quantities we can uniquely recovers the primitives $(u_{\bar{k}j}, \psi_{\bar{k}j})$. However, in order to conduct a counterfactual analysis, such as considering the effect of an exogenous change in worker-firm specific productivities $\gamma_{\bar{k}j}$, the unemployment benefit $w_{\bar{k}0}$, or workers preferences for firms amenities u_{kj} on the equilibrium wages and matches of workers across firms $(w_{\bar{k}j}, s_{\bar{k}j})$ we must analyze the inverse identification problem. More precisely, do knowing the parameters $(u_{\bar{k}j}, \psi_{\bar{k}j})_{(k,j)\in(\mathcal{K}\times\mathcal{J})}$, and the unemployment benefits, $(w_{\bar{k}0})_{\bar{k}\in\mathcal{K}}$ allow us to uniquely recover the aggregate wages and workers matching distributions across firms, i.e. $(w_{\bar{k}j}, w_{\bar{k}0}, s_{\bar{k}j})_{(\bar{k},j)\in(\mathcal{K}\times\mathcal{J})}$? In general, in a model with a finite number of firms there is an important presence of strategic interactions across firms which may lead to multiple equilibria and therefore make the counterfactual analysis very complicated. Card et al (2018) pointed out this issue in the following term:

"By assuming that the number of employers is very large, we have adopted a partial equilibrium framework with no strategic interactions between employers. With a finite number of firms, a shock to one firm's productivity will affect the equilibrium employment and wages of competitor firms. Staiger, Spetz, and Phibbs (2010) provide compelling evidence of such responses in the market for nurses. As in the oligopoly literature, analysis of a finite employer model with strong strategic dependence may be complicated by the presence of multiple equilibria, which requires different methods for estimation (e.g., Ciliberto and Tamer 2009) but may also yield interesting policy implications."

The next proposition shows that in our model with a finite number of firms and linear production functions, the inverse problem has a unique solution, and thus, we are therefore in the presence of a unique equilibrium.

Proposition 4 (Inverse Identification Problem)**.** *Consider Assumptions 3, 4, and 5. For a fixed set of parameters* $(u_{kj}, \psi_{\bar{k}j})_{(k,j)\in(\mathcal{K}\times\mathcal{J})}$, *and unemployment benefits* $(w_{\bar{k}0})_{\bar{k}\in\mathcal{K}}$ *whenever Assumption 9 holds there always exists a unique endogenous aggregate wage distribution and matching pattern* $(w_{\bar{k}j}, s_{\bar{k}j})_{(\bar{k},j)\in(\mathcal{K}\times\mathcal{J})}$ *that solves the equilibrium model characterization, i.e. eqs (3.13, 3.14) and the population constraints eq(2.11).*

In fact, for a fixed set of primitives, unemployment benefits and workers population supply, equations (3.13, 3.14) allow us to uniquely recover $(w_{\bar{k}j}, s_{\bar{k}j})$ up to the endogenous unknown vector of unemployed $(s_{\bar{k}0})_{\bar{k}\in\mathcal{K}}$. So, the main question to verify is whether there is a unique positive vector of unemployed shares $(s_{\bar{k}0} > 0)_{\bar{k}\in\mathcal{K}}$ that solves the following system of population constraints:

$$s_{\bar{k}0} + J_{\bar{k}} + \sum_{j\in\mathcal{J}_{\bar{k}}} \left\{ \frac{1}{2}\psi_{\bar{k}j}\frac{u_{\bar{k}j}}{w_{\bar{k}0}} s_{\bar{k}0} - \sqrt{\left(\frac{1}{2}\psi_{\bar{k}j}\frac{u_{\bar{k}j}}{w_{\bar{k}0}} s_{\bar{k}0}\right)^2 + 1} \right\} = 1 \text{ for } \bar{k} \in \mathcal{K}, \tag{5.7}$$

and the answer is yes as a corollary of Lemma 1. The key for uniqueness to hold is the fact that firms have linear production functions —workers are perfect substitutes. If we consider that workers are imperfect substitutes the uniqueness result is no longer guaranteed to hold and we could be in presence of multiple equilibria. Proposition 3 and 4 shows that in a single market any observed aggregate distribution of wages and matching pattern can be rationalized exactly by a many-to-one matching model under the Logit economy assumption.

The next section is a work-in-progress, comments are more than welcome.

5.2. **When firms set wages using workers' unobservable types.** As we previously discussed, Assumption 9 can be quite restrictive. However, we see that under this assumption the model fits exactly any aggregate data on a single cross-sectional market. If we relax Assumption 9, and allow firms to set wages based on characteristics that are unobservables to the econometrician, i.e. $\tilde{k}$; then without additional information, we lose the (point) identification results in Proposition 3 and Corollary 2. To try to recover identification, we consider that we have access to a panel data, where workers and firms are observed in multi-periods, i.e. $t = 1, ..., T$. We propose to recover $\tilde{k}$ using the *kmeans*'s distribution-based approach discussed in Bonhomme, Lamadon and Manresa (2017, BLM17). The main goal of the kmeans classification approach is to group together workers with similar latent attributes which are unobserved to the econometrician. As discussed in BLM17 the kmeans algorithm is a popular tool often used and studied in machine learning and computer science.[24] Below, we describe a heuristic two-step approach that would allow us to consistently estimates the key structural parameters of the model. For now, we provide the intuition behind our identification strategy. Since Choo and Siow (2006), the identification of structural parameters in the matching literature has mainly used *only* the DMF to recover structural parameters and this

[24]Please refer to BLM17 for a more detailed description of the method.

almost always has relied on Assumption 9. The main rationale is that in the marriage literature the transfers (wages in our case) are not observed, so the equation that links the transfers to the structural parameters cannot be used for identification. This makes the model highly under-identified without additional restrictions which is why Assumption 9 is important in that framework. On the other side, in the empirical labor literature, since AKM the identification of the structural parameters makes use of only the EWF. The rationale is that AKM and Bonhomme, Lamadon and Manresa (2018) are essentially reduced-form equations without a structural model that would explain the link between the DMFs and the structural parameters. But notice that even if the DMFs are not explicitly characterized in their cases, their identification of the firms fixed-effects requires movers, which in some sense contain information on the shares, and hence the DMFs. Here, we propose to explicitly use the mapping —between (EWFs & DMFs), and structural parameters— induced by the model we entertain to recover key structural parameters. We consider that observing wages may allow one to recover worker unobserved types, and then once estimated we can recover identification of the structural parameters as in Proposition 3, and Corollary 2. In other words, observing wages in multi-periods may allow us to relax Assumption 9. Notice that Fox (2018) considers a model where the transfers are observed but does not take advantage of this additional information for identification, using only the DMF for identification under Assumption 9. Our current approach could be adapted to Fox's context in order to relax Assumption 9.

Recovering $\tilde{k}$ using kmeans' distribution-based approach. If the researcher has a good understanding of the observed summary statistics that could be informative to recover $\tilde{k}$, the researcher could proceed to his classification using only those summary statistics, however such a classification is model-specific. A general approach which does not rely on researcher model specification is to make use of the entire empirical distribution of the data, in order to capture all the relevant heterogeneity in the classification step. Let $W_{it} = (Y_{it}, S_{it}, \bar{X}'_{it})'$ a vector of worker i outcomes and covariates for $i = 1, ..., N$, where Y_{it} denotes the observed income (included the unemployment benefit for unemployed workers), $S_{it} \in \{0, 1, ..., J\}$ denotes the firm choice (firm 0, refers to unemployment), and $\bar{X}_{it}$ denotes the vector of workers exogenous covariates observed by econometricians, and all this for every period $t \in \{1, ..., T\}$. Let $\hat{F}_i(w) = \frac{1}{T}\sum_{t=1}^{T} 1\{W_{it} \leq w\}$ denotes the empirical CDF of W_{it}. We follow essentially BLM17 and consider the following minimization problem:

$$(\hat{\tilde{k}}_i, ..., \hat{\tilde{k}}_N, \hat{H}) = \underset{(H, \tilde{k}_1, ..., \tilde{k}_N)}{argmin} \sum_{i=1}^{N} \int \left(\hat{F}_i(w) - H_{\hat{\tilde{k}}_i}(w)\right)^2 d\mu(w) \tag{5.8}$$

with $H(w) = (H_1(w), ..., H_{\tilde{K}}(w))'$ where $H_{\tilde{k}}(w)$ for $\tilde{k} \in \{1, ..., \tilde{K}\}$ are CDFs, $(\tilde{k}_1, ..., \tilde{k}_N) \in \{1, ..., \tilde{K}\}^N$ are partition of $\{1, ..., N\}$ into at most $\tilde{K}$ groups. $\tilde{K}$ is the cardinality of unobservable type to econometricians, it is ex-ante fixed by the researcher but BLM17 proposes a data-driven way to obtain $\tilde{K}$ that we follow here. The kmeans procedure will therefore provides a partition of the N workers into $\tilde{K}$ groups, such that for each worker i, we have $\hat{\tilde{k}}_i = l$ for $l = 1, ..., \tilde{K}$.

Identification using estimated unobserved types $\hat{\tilde{k}}$. Now, for each worker i, the econometrician "knowns" (or has a consistent estimator) to which group $\hat{k} = (\hat{\tilde{k}}, \bar{k})$ she belongs to. We can therefore construct the aggregate equilibrium wages and shares, $w_{\hat{k}jt}, s_{\hat{k}jt}$ which are, respectively, the average wages and the proportion of worker of type $\hat{k}$ workers, working at firm j at the period $t \in \{1, ..., T\}$. So, whenever Assumptions 3, and 4 hold and under the assumption that $\tilde{k}$ are consistently estimated using the kmeans approach in the first step, we have the following identification result:

$$
\begin{aligned}
u_{\hat{k}jt} &= \frac{w_{\hat{k}0}}{w_{\hat{k}jt}} \frac{s_{\hat{k}jt}}{(1 - \sum_{j \in \mathcal{J}_{\hat{k}}} s_{\hat{k}jt})}, \quad \hat{k} \in \mathcal{C}^j, j \in \mathcal{J} \\
\psi_{\hat{k}jt} &= w_{\hat{k}jt} \frac{2 - s_{\hat{k}jt}}{1 - s_{\hat{k}jt}}, \quad \hat{k} \in \mathcal{C}^j, j \in \mathcal{J}
\end{aligned}
$$

$$
\phi_{\hat{k},j,\hat{k}',j'} = \ln \frac{w_{\hat{k}j} w_{\hat{k}'j'}}{w_{\hat{k}jt'} w_{\hat{k}'j}} + \ln \Big(\frac{2 - s_{\hat{k}jt}}{1 - s_{\hat{k}jt}} \Big) \Big(\frac{2 - s_{\hat{k}'j'}}{1 - s_{\hat{k}'j'}} \Big) \Big(\frac{1 - s_{\hat{k}'j}}{2 - s_{\hat{k}'j}} \Big) \Big(\frac{1 - s_{\hat{k}jt'}}{2 - s_{\hat{k}jt'}} \Big), \ \hat{k}, \hat{k}' \in \mathcal{C}^j \cap \mathcal{C}^{j'}.
$$

6. Empirical Illustration

In this section we apply our identification strategy using a detailed employer-employee matched dataset from Denmark. For now we will focus on identifying the case where wages are based only on workers' observable types. This allows us to easily identify the structural parameters of interest as well as our measure of ALC, which we can then use to test for assortative matching between workers and firms.

6.1. **Data Description.** In order to conduct our empirical analysis, we merge several administrative data registers produced by Statistics Denmark. The IDA register provides annual worker-level information (measured in November of that year) on wages, age, education and other demographics for the entire population of Denmark. Our firm-level data comes from the FIRE and FIRM registers, which provide firm-level measures of revenues, employment, intermediate material expenditure, labor costs, capital stock and industry. The employee-employer linkages in FIDA allow us to merge these data. Additionally, we obtain measures of unemployment income and benefits for every individual from the IND register. Our worker sample consists of all individuals working at their primary November job in the private sector, thus excluding employers as well as public sector employees. This leaves us with just over 2,550,000 individual observations in 2005.

We assign workers into types ($\bar{k}$) where each type is a tuple of sex, age and education. We set three age bins: {20-35, 36-50, 51-65} and four education bins: {High school or less, Vocational education, 2-4 years of university, 5+ years of university}. This results in 24 types based on observable characteristics. We then use the worker-firm linkages to calculate employment shares $s_{\bar{k}j}$ for each firm j of each worker type $\bar{k}$.

Our analysis is done using a single annual cross-section for the year 2005 which provides us with over 130,000 firms and 510,000 j-k pairs. We define the wage w_{ij} for a worker i in firm j as the annualized hourly income for that worker at that firm *had they worked at the firm for the entire* *year as a full time employee.* In particular, we scale each worker's hourly wage by a constant – the average number of full time hours worked by full-time workers in Denmark. This controls for differences in hours worked in our measure of annual wage. Similarly, the unemployment benefits w_{i0} for a particular worker i are defined as the total annualized unemployment income[25] *had that worker been unemployed for the entire year.* We then aggregate these up to the $\bar{k}$ type level by calculating the mean income $w_{\bar{k}j}$ for each worker type at each firm, and the mean unemployment income $w_{\bar{k}0}$ for each worker type in the economy. This provides all of the information required to identify our key structural parameters of interest: $u_{\bar{k}j}$ and $\psi_{\bar{k}j}$.

6.2. **Results.** We have several goals with our empirical analysis. We want to understand the firm and market characteristics which drive differences in non-pecuniary benefits and wage setting. Additionally our framework allows us to speak to the degree of assortative matching present in the data. We perform several exercises along these lines.

As mentioned above, not every firm need employ every type of worker. Indeed, depending on the granularity of $\bar{k}$, it may be highly unusual for even very large firms to employ every worker type. This can be seen in the Danish data. Table 1 shows the share of firms which employ each of the 24 types of workers we have defined in the data. The worker group most employed is middle-aged men with vocational educations (34% of firms), followed by young men with vocational educations or less. The least employed worker type is women with 5 or more years of college, which includes masters degrees and phds. Figure 1 shows that while most firms employ 3 or fewer of the worker types we define, there is a significant right tail, with over 400 firms (0.36% of all firms) employing all 24 worker types in 2005.

One of our key parameters of interest is our measure of non-pecuniary amenities or firm-type preferences. Given the firm-type calculation of non-pecuniary benefits $u_{\bar{k}j}$, we can calculate u_j^{np} which represents the normalized aggregate preferences over all workers working at firm j. Figure 2 plots the density of u_j^{np} (re-scaled by 1000), truncating the top and bottom percentiles. The distribution is highly right skewed. Most firms are clustered towards 0, with a few firms having relatively much higher amenity values. Table 2 shows moments of the re-scaled amenity distribution by industry. The overall mean amenity value is 0.31 with a standard deviation of 2.78. The "professional, scientific and technical" industry has both the highest mean amenity level (0.37) as well as the largest dispersion (4.20). Amenity levels in the "ICT" and "manufacturing" industries are also above the overall average. The "other services", "construction" and "finance" industries have the lowest amenity levels, at 0.17, 0.19 and 0.2 respectively. Since we are ultimately interested in what firm-level characteristics may drive this dispersion in amenities or preferences over firms, we run a simple OLS regression of log firm amenities on logged firm characteristics, which we report in table

[25]Our measure of unemployment income includes cash assistance, unemployment benefits, leave benefits and other assistance benefits, but similar to our measure of wages does not include long-term sickness or pension benefits.

3. The first four columns show regressions of the following form:

$$\log u_j^{np} = Z_j\beta + \gamma_i + \varepsilon_j \tag{6.1}$$

where γ_i is an industry fixed effect and X_j is a vector of firm level characteristics including revenue (R), employment in full-time equivalents (L), capital stock (K), expenditure on intermediate goods and services (M), total labor cost (W), value-added per worker (VA/L), average daily hours worked per worker (H/L), and the number of worker types employed by the firm ($|C^j|$). The last four columns run a similar regression, but at the firm-type level:

$$\log u_{\bar{k}j} = Z_j\beta + \gamma_i + \gamma_k + \varepsilon_j \tag{6.2}$$

where we can now add a worker type fixed effect, γ_k, and the LHS variable is the amenity value for firm j and worker type k. These regressions allow us to look at correlations between firm characteristics and amenity values, though don't necessarily speak to the causal relationship. At the firm-level, we find that given employment and labor costs, firms which are more productive and employ more worker types tend to have higher amenity values. Amenities tend to be decreasing in capital stock, however. Employment is strongly positively correlated with amenities, while labor cost is strongly negatively correlated. When ignoring the other firm-level characteristics, amenities appear to be positively correlated with the average number of hours worked by workers at the firm. However, the sign flips in the full regression, indicating that conditional on wages, revenue and other firm characteristics, firms which have longer working hours have significantly lower amenity values. Moving to the firm-type amenity regressions, we find much higher positive (negative) correlations with employment (labor costs). Controlling for both industry and worker type fixed effects flips the sign on log capital, suggesting a positive relationship between capital stock and amenities.

We can also look at how workers of each type value the amenities of their employers on average. This can give us an idea of how the amenities offered by firms differ by worker type. Table 4 shows the average (re-scaled) value of firm-type amenities $u_{\bar{k}j}$ by worker type $\bar{k}$. There are a few striking results. First, through the lens of the model, female workers must value the amenities of the firms where they work much more than men on average in order to rationalize the differences in wages and employment shares. Second, workers with some college or a bachelor's degree tend to work at firms over which they have higher preferences/amenities than workers with more or less education. The average amenity value for middle aged men with 2-4 years of college education (2.57) is an order of magnitude greater than middle aged workers with a vocational education (0.24), and 5 times higher than middle aged men with postgraduate degrees (0.51). Low education middle aged men also tend to work at firms with lower amenity values than their younger and older counterparts, while the relationship is flipped for those with at least a few years of college education.

Finally, we apply our identification strategy to identify our measure of aggregate local complementarity (ALC) in the data. Given a ranking over workers and firms, we are able to point-identify the ALC ($\phi_{\bar{k},j,\bar{k}',j'}$) for each set of $\bar{k}, j$ and $\bar{k}', j'$. In practice, we partition firms into deciles $\bar{j}$ based on a ranking over j and then calculate $\phi_{\bar{k},\bar{j},\bar{k}',\bar{j}'}$ for each $\bar{k}, \bar{j}$ where $\bar{k}' \equiv \bar{k} - 1$ and $\bar{j}' \equiv \bar{j} - 1$. We perform these calculations for each combination of age and sex, where we look at worker sorting over education. We also do this for two different rankings over firms: their estimated amenity value u_j^{np}

and their value added per worker (VA/L) as a proxy for firm productivity. We report the results for middle-aged male workers in tables 5 and 6. We also report bootstrapped 95% confidence intervals below each point estimate in brackets[26]. Table 5 shows the point estimates of the ALC between education and value-added per worker for the discrete decile and education categories defined above. A weak majority are positive (16 out of 27), with more evidence for positive assortative matching for workers with a college education. However few of these estimates are significantly different from zero. Table 6 shows a similar pattern for firm amenities, with very highly educated workers appearing to be positively sorted into high amenity firms. However, the ALC estimates are uniformly negative for workers with only some college. This may suggest that the ranking over education, in particular between vocational education and some college, is not particularly strong. Our point estimates for the firm amenity rankings are much more precise relative to the value-added per worker ranking, with most point estimates being significantly different from zero. Overall we cannot yet conclude at this point that the data exhibits either positive or negative assortative matching. We are currently working on a more formal test which we can apply to the data.

Work-in-progress.

7. Conclusion

This paper developed a theoretical framework consisting of a two-sided matching model of the labor market. We showed how to identify preferences for firm amenities using a revealed preference approach and additionally how to identify the pattern of complementarities between workers and firms. The most obvious extension of this framework is to introduce dynamics, in particular, mobility of workers across firms, which may also faciliate identification of structural parameters. Although outside of the scope of this paper, the next step would be to consider an empirical application using matched employee-employer earnings data.

[26]See appendix for a description of our bootstrap methodology.

TABLE 1. Share of all firms employing each worker type

	Males			
Age	High School or Less	Vocational	2-4 Years of College	5+ Years of College
20-35	0.30	0.31	0.13	0.05
36-50	0.24	0.34	0.15	0.07
51-65	0.18	0.28	0.11	0.05

	Females			
Age	High School or Less	Vocational	2-4 Years of College	5+ Years of College
20-35	0.22	0.20	0.18	0.05
36-50	0.19	0.24	0.13	0.05
51-65	0.15	0.18	0.08	0.02

FIGURE 1. Share of firms by number of worker types employed

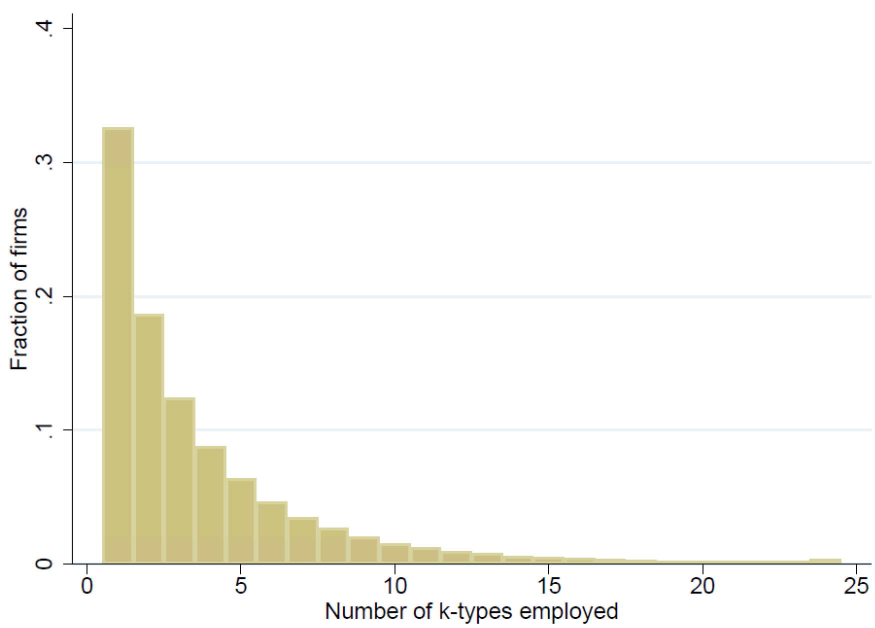


Table 2. Moments of Firm Amenity Distribution by Industry

	Firm Amenities (u_j^{np})	
Industry	Mean	Standard Deviation
ALL	0.31	2.78
Mining	0.24	0.37
Manufacturing	0.32	1.01
Construction	0.19	0.26
Wholesale and Resale	0.22	0.60
Transportation	0.22	1.47
Accommodation and Food Services	0.30	0.52
Information and Communications	0.34	1.07
Finance and Insurance	0.20	0.32
Real Estate	0.21	0.52
Professional, Scientific, Technical	0.37	4.20
Administrative and Support Services	0.31	1.00
Other Services	0.17	0.21

Note: Table 2 shows the industry-level moments of u_j^{np}, where u_j^{np} has been rescaled by 1000. Only industries with at least 100 firms in 2005 have been reported here.

Table 3. Regressions of Firm Amenities on Firm-Level Characteristics

	Firm Amenities ($\log u_j^{np}$)				Firm-Type Amenities ($\log u_{\bar{k}j}$)			
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log L	0.511***	0.560***		0.648***	0.718***	0.726***	0.827***	0.861***
	(0.009)	(0.009)		(0.0.009)	(0.014)	(0.015)	(0.014)	(0.014)
Log K	-0.030***	-0.025***		-0.026***	-0.018***	-0.012***	0.009***	0.008***
	(0.002)	(0.002)		(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Log M	-0.018***	0.022***		0.024***	-0.015***	0.002	-0.011***	-0.008***
	(0.004)	(0.004)		(0.004)	(0.005)	(0.005)	(0.005)	(0.004)
Log R	0.003	0.018*		0.010	-0.020**	-0.024**	0.018*	0.014
	(0.009)	(0.009)		(0.009)	(0.009)	(0.010)	(0.009)	(0.009)
Log VA/L	0.077***	0.046***		0.047***	0.107***	0.076***	0.072***	0.076***
	(0.007)	(0.007)		(0.0.007)	(0.008)	(0.008)	(0.007)	(0.007)
Log W	-0.339***	-0.407***		-0.415***	-0.399***	-0.441***	-0.532***	-0.538***
	(0.007)	(0.007)		(0.007)	(0.010)	(0.011)	(0.010)	(0.010)
$\lvert C^j \rvert$	0.074***	0.065***		0.047***	0.027***	0.019***	0.007***	0.001
	(0.001)	(0.001)		(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Log H/L			0.133***	-0.411***				-0.196***
			(0.007)	(0.007)				(0.006)
Industry FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
K-type FE	No	No	No	No	No	No	Yes	Yes
R^2	0.317	0.349	0.029	0.371	0.237	0.244	0.531	0.536
Observations	83,379	83,379	83,379	83,379	346,090	346,090	346,090	346,090

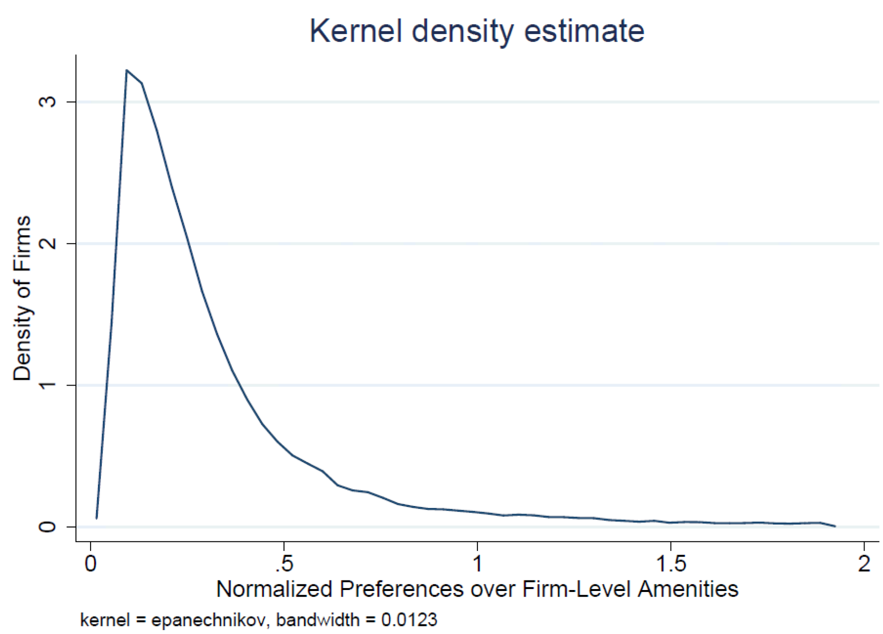
Note: All variables are in logs except $|C^j|$, which is the count of worker types employed by firm j. Variables L, K, M, R, VA/L, W, and H/L are (respectively) employment in full time equivalents, capital stock, material expenditures, revenue, value-added per worker, total labor cost, and average daily hours worked per worker.

Table 4. Mean amenity value for each worker type

	Males			
Age	High School or Less	Vocational	2-4 Years of College	5+ Years of College
20-35	0.48	0.27	1.52	0.49
36-50	0.33	0.24	2.57	0.51
51-65	0.45	0.28	1.16	0.49

	Females			
Age	High School or Less	Vocational	2-4 Years of College	5+ Years of College
20-35	0.60	0.44	2.69	0.78
36-50	0.26	0.46	3.28	0.61
51-65	0.42	0.29	1.78	0.71

Figure 2. Distribution of u_j^{np}



Note: Figure 2 shows the distribution of u_j^{np}, where u_j^{np} has been rescaled by 1000.

Table 5. Aggregate Local Complementarity ()$\phi_{\bar{k},\bar{j},\bar{k}',\bar{j}'}$) between Education and Value Added per Worker

	Education Group for 36-50 Year Old Males			
Decile of Firm VA/L	High School or less	Vocational	2-4 Years of College	5+ Years of College
1	-	-	-	-
2	-	0	-0.003	0.040
		(-0.017, 0.015)	(-0.041, 0.033)	(-0.023, 0.107)
3	-	-0.019	0.032	0.007
		(-0.030, 0.003)	(0.006, 0.054)	(-0.051, 0.062)
4	-	-0.028	0.024	0.045
		(-0.045, -0.009)	(0.004, 0.051)	(-0.030, 0.074)
5	-	-0.012	0.014	-0.051
		(-0.031, 0.010)	(-0.012, 0.036)	(-0.086, 0.010)
6	-	-0.028	0.035	0.002
		(-0.059, -0.001)	(0.005, 0.060)	(-0.020, 0.044)
7	-	0.014	-0.014	0.119
		(-0.038, 0.052)	(-0.046, 0.021)	(0.057, 0.156)
8	-	0.045	-0.039	-0.123
		(-0.008, 0.099)	(-0.094, 0.003)	(-0.173, -0.026)
9	-	0.014	0.012	0.042
		(-0.044, 0.073)	(-0.035, 0.074)	(-0.037, 0.099)
10	-	-0.063	0.013	-0.050
		(-0.115, -0.007)	(-0.051, 0.064)	(-0.120, 0.010)

Note: The bootstrapped 95% confidence intervals are in brackets below the ALC point estimates. Each estimate is with respect to the next firm rank decile and education level down, which is why there are no estimates for the lowest firm decile or education level.

Table 6. Aggregate Local Complementarity ($\phi_{\bar{k},\bar{j},\bar{k}',\bar{j}'}$) between Education and Firm Amenities (u_j^{np})

	Education Group for 36-50 Year Old Males			
Decile of u_j^{np}	High School or less	Vocational	2-4 Years of College	5+ Years of College
1	-	-	-	-
2	-	0.083	-0.012	-0.396
		(0.041, 0.132)	(-0.197, 0.163)	(-0.604, -0.186)
3	-	0.048	-0.468	0.244
		(0.005, 0.069)	(-0.632, -0.293)	(0.032, 0.432)
4	-	-0.037	-0.259	0.204
		(-0.062, 0.005)	(-0.396, -0.090)	(0.033, 0.365)
5	-	-0.004	-0.392	0.333
		(-0.031, 0.024)	(-0.488, -0.294)	(0.186, 0.434)
6	-	0.008	-0.264	0.244
		(-0.019, 0.033)	(-0.311, -0.211)	(0.163, 0.326)
7	-	-0.011	-0.208	0.146
		(-0.029, 0.022)	(-0.238, -0.166)	(0.091, 0.219)
8	-	0.024	-0.131	0.125
		(-0.004 0.042)	(-0.149, -0.010)	(0.080, 0.171)
9	-	-0.008	-0.067	0.063
		(-0.022, 0.014)	(-0.086, -0.048)	(0.014, 0.097)
10	-	-0.002	-0.062	0.064
		(-0.023, 0.011)	(-0.070, -0.038)	(0.018, 0.082)

Note: The bootstrapped 95% confidence intervals are in brackets below the ALC point estimates. Each estimate is with respect to the next firm rank decile and education level down, which is why there are no estimates for the lowest firm decile or education level.

Appendix A. Detailed derivations of formula in the main text.

A.1. **Four-parameter Beta distribution.** Manning (2011) has shown that the equilibrium wage in a wage-posting model is isomorphic to the wage in an ex post wage-bargaining framework. To show this in our context, let's assume that the random variable $e^{\epsilon_{k_i0}-\epsilon_{k_i1}}$ follows the four-parameter Beta distribution, i.e. $e^{\epsilon_{k_i0}-\epsilon_{k_i1}} \sim B(\beta, 1, a, b)$ with $\beta > 0$ and $0 < a < b$, where β, and 1 are shape parameters and $[a, b]$ denotes the support of $e^{\epsilon_{k_i0}-\epsilon_{k_i1}}$, i.e. $Supp(e^{\epsilon_{k_i0}-\epsilon_{k_i1}}) = [a, b]$. After some derivations, we obtain the equilibrium wage and distribution matching function (EWF & DMF) as function of the primitives only:

$$w_{k1} = \frac{\beta}{1+\beta}(\lambda_1\theta_1\gamma_{k1}) + \frac{1}{1+\beta}\underline{w}^R_{k1}, \text{ for } \underline{w}^R_{k1} \leq \lambda_1\theta_1\gamma_{k1} \leq \overline{w}^R_{k1} + \frac{1}{\beta}(\overline{w}^R_{k1} - \underline{w}^R_{k1}), \tag{A.1}$$

and

$$s_{k1} = \begin{cases} 0 & \text{if } \lambda_1\theta_1\gamma_{k1} < \underline{w}^R_{k1}, \\ \left(\frac{\beta}{1+\beta}\right)^\beta \left(\frac{\lambda_1\theta_1\gamma_{k1} - \underline{w}^R_{k1}}{\overline{w}^R_{k1} - \underline{w}^R_{k1}}\right)^\beta & \text{if } \underline{w}^R_{k1} \leq \lambda_1\theta_1\gamma_{k1} \leq \overline{w}^R_{k1} + \frac{1}{\beta}(\overline{w}^R_{k1} - \underline{w}^R_{k1}), \\ 1 & \text{if } \lambda_1\theta_1\gamma_{k1} > \overline{w}^R_{k1} + \frac{1}{\beta}(\overline{w}^R_{k1} - \underline{w}^R_{k1}). \end{cases} \tag{A.2}$$

where $\underline{w}^R_{k1} \equiv \frac{w_{k0}}{u_{k1}} a = \min\left\{\frac{w_{k0}}{u_{k1}} e^{\epsilon_{k_i0}-\epsilon_{k_i1}} : e^{\epsilon_{k_i0}-\epsilon_{k_i1}} \in [a, b]\right\}$ and similarly $\overline{w}^R_{k1} \equiv \frac{w_{k0}}{u_{k1}} b$.

Intuitively, the EWF depends on the shape parameter of the distribution of the shocks and the minimum reservation wage. The shape parameter governs the slope of the labor supply function. On the other hand, the minimum reservation wage pins down the level of the labor supply function. A higher value increases the EWF. Thus, both the shape parameter and the level of reservation wages are relevant for the firm's optimal wage decision and are what drive the labor supply elasticity. To connect this equation to the wage bargaining solution, recall that $\lambda_1 > 0$ is the marginal cost of output. At the profit maximizing level of output, λ_1 is equal to marginal revenue product. In the case of perfectly competitive firms in the output market, $\lambda_1 = P_1$ where P_1 is the market price. If there is no heterogeneity in workers skills, we can set $\theta_1\gamma_{k1} = 1$. In this case, we recover the well known equilibrium equation obtained in the ex post wage-bargaining framework:

$$w_{k1} = \frac{\beta}{1+\beta}P_1 + \frac{1}{1+\beta}\underline{w}^R_{k1}, \tag{A.3}$$

where bargaining power is entirely captured by the shape parameter β.[27] This shows that the wage-posting model we consider here can rationalize the equilibrium wage predicted by the ex post wage-bargaining model. Equation (A.1) can be viewed as a generalization of the ex post wage-bargaining equilibrium when there are heterogenous workers and the product market is not perfectly competitive.[28]

[27]The ex post wage-bargaining is a model where the wage is split after the worker and employer have been matched, according to some sharing rule, such that an asymmetric Nash bargain.

[28]For instance, in the presence of downward-sloping firm-specific product demand $\lambda_j(\eta) = (\eta - \frac{1}{\eta})P_j(Y^j)^{-\frac{1}{\eta}}$, for $\eta > 0$.

A.2. **Beckerian wage discrimination.** An early attempt to rationalize wage discrimination in a competitive market was made by Becker (1957). Becker argued that some employers dislike employing women/black/minorities and that this can be modelled by allowing for the presence of a distaste parameter in the employers' utility function. To fix the ideas, let's revisit Becker's discrimination theory within our present framework. Firm j's utility can be expressed as follows:

$$V_j = \tilde{\gamma}_{\tilde{k}j}\Big[\mu_{kj} + \mu_{k'j}\Big] - d_{k'j}\mu_{k'j} - \Big[w_{kj}\mu_{kj} + w_{k'j}\mu_{k'j}\Big] \tag{A.4}$$

$$= \Big[\underbrace{\tilde{\gamma}_{\tilde{k}j}}_{\gamma_{(m,\tilde{k})j}}\ \mu_{kj} + \underbrace{(\tilde{\gamma}_{\tilde{k}j} - d_{k'j})}_{\gamma_{(f,\tilde{k})j}}\mu_{k'j}\Big] - \Big[w_{kj}\mu_{kj} + w_{k'j}\mu_{k'j}\Big]. \tag{A.5}$$

In Becker's model, both types of worker have the same productivity $\tilde{\gamma}_{\tilde{k}j}$ but each women/black hire by the firm generates disutility $d_{k'j} \geq 0$. Solving the model as if we were in a competitive market leads to the BWG derived earlier in eq (3.5). In Becker's model, the marginal firms with the lowest distaste for women/black determine the women/black wages in equilibrium. Since firms with the high distaste sacrifice profits by discriminating, we may expect such firms to be driven away from the market in the long run. Therefore the BWG will disappear in long run. This was indeed the famous Arrow's (1972) criticism about Becker's discrimination model. However, as discussed in Charles and Guryan (2008) some empirical studies seems to suggest some important presence of distaste and prejudicial feelings for minorities in U.S. Arrow and other papers have argued that BWG can survive in the long run with imperfect competition as we consider here. Before analyzing the RWG, notice that in our general framework without explicitly modelling the disutility we can rationalize the Becker's by allowing for $\gamma_{(m,\tilde{k})j} \neq \gamma_{(f,\tilde{k})j}$; in other words, as a result of discrimination or distaste, a firm can evaluate the marginal contributions of two different type of workers differently even if they have the same level of skill.

A.3. **Nested Logit Economy.** Let's partition the J firms into G classes, the g^{th} class contains N_g firms. g_j denotes the firm j belonging to the class g. The errors ϵ_{ig_j} are assumed to be correlated within classes, i.e. $\rho_g = \sqrt{1 - corr(\epsilon_{ig_j}, \epsilon_{ig_l})} \in (0,1)$ for $j \neq l$. In this nested Logit framework we have that $G_{k\cdot}(v_{k\cdot}) = \ln\Big\{e^{v_{k0}} + \sum_{g=1}^{G}\Big(\sum_{j=1}^{N_g} e^{v_{kg_j}/\rho_g}\Big)^{\rho_g}\Big\}$. After derivations, one can show that $\frac{\mu_{kg_j}}{\mu_{k0}} = e^{v_{kg_j}/\rho_g - v_{k0}}\Big(\sum_{l=1}^{N_g} e^{v_{kg_l}/\rho_g}\Big)^{\rho_g - 1}$, $\frac{\mu_{kg}}{\mu_{k0}} = e^{-v_{k0}}\Big(\sum_{l=1}^{N_g} e^{v_{kg_l}/\rho_g}\Big)^{\rho_g}$ where $\mu_{kg} = \sum_{j=1}^{N_g}\mu_{kg_j}$ and then the labor supply elasticity in this Nested Logit economy can be written as follows:

$$\mathcal{E}_{kg_j} = \sigma_g + (1 - \sigma_g)\frac{s_{kg_j}}{s_{kg}} - s_{kg_j} \tag{A.6}$$

where $\sigma_g \equiv 1/\rho_g$, $s_{kg_j} = \frac{\mu_{kg_j}}{m_k}$, $s_{kg} = \frac{\mu_{kg}}{m_k}$. Notice that when we have perfect correlation of the errors within classes, in others terms firms are perfect substitute within nests $\rho_g = 0 \Rightarrow \sigma_g \to \infty$ then $\frac{\mathcal{E}_{kg_j}}{1+\mathcal{E}_{kg_j}} \to 1$, therefore the EWF becomes

$$w_{kg_j} = \lambda_{g_j} F_k^{g_j}(\mu_{\cdot g_j}), \text{ for all } k \in \mathcal{C}^{g_j} \tag{A.7}$$

which is the EWF in the perfect competition case. Now let consider that $\sigma_g < \infty$, we can derive the following quasi-supply:

$$s_{kg_j} = s_{k0}^{\sigma_g} s_{kg}^{(1-\sigma_g)} e^{\sigma_g(v_{kg_j} - v_{k0})}. \tag{A.8}$$

For the sake of simplicity, let's consider Assumption 3. Then, by using the above equations and eq (2.8) we can show that, the EWF in the nested Logit economy is the lowest non-negative solution of this non-linear equation:

$$A(\sigma_g, s_{k0}, s_{kg}) w_{kg_j}^{1+\sigma_g} - \psi_{kg_j} A(\sigma_g, s_{k0}, s_{kg}) w_{kg_j}^{\sigma_g} + (1+\sigma_g) w_{kg_j} - \psi_{kg_j}\sigma_g = 0, \tag{A.9}$$

where $A(\sigma_g, s_{k0}, s_{kg}) \equiv \Big\{(1-\sigma_g) - s_{kg}\Big\}\big(\frac{s_{k0}}{s_{kg}}\big)^{\sigma_g}\big(\frac{u_{kj}}{w_{k0}}\big)^{\sigma_g}$.

It is worth-noting that the above equations can be solved numerically. In the light of the above equations, we can see that, in this nested Logit Economy, the resulting log EWF is likely to be highly nonlinear in unobserved worker and firm heterogeneity. Therefore, AKM or Bonhomme et al (2017) proposed reduced form wage equations could be an over-simplification of the actual equilibrium wage equation, which therefore may generates significant biases. Using an adequate monte-carlo exercise the above equation can allow to judge how biases could be an AKM estimates when we are in a Nested Logit Economy.

A.4. **Derivations of the EWFs.** Here, we will derive the structural equilibrium wage functions (EWFs) and distributional worker-firm matching functions (DMFs) related to the considered production function, i.e. $F^j(\mu_{\cdot j}) = \theta_j\Big(\sum_{k=1}^K \gamma_{kj}\mu_{kj}\Big)^{\alpha_j}$ when the choice probabilities of workers are given by:

$$\begin{aligned} s_{kj}(v_{k\cdot}) &= \chi_0(k)\Big(\frac{w_{kj}u_{kj}}{w_{k0}u_{k0}}\Big)^{\beta_j}, \qquad (A.10)\\ s_{k0}(v_{k\cdot}) &= 1 - \sum_{j\in\mathcal{J}} s_{kj}(v_{k\cdot}). \qquad (A.11)\end{aligned}$$

Using this production function, $\frac{\partial F^j(\mu_{\cdot j})}{\partial \mu_{kj}} = \theta_j\alpha_j\gamma_{kj}\Big(L_j\Big)^{\alpha_j - 1}$, then the optimal wage equation (2.8) simplifies to:

$$w_{kj} = \lambda_j\theta_j\frac{\alpha_j\beta_j}{1+\beta_j}\gamma_{kj}\Big(L_j\Big)^{\alpha_j-1} \equiv A_j\gamma_{kj}\Big(L_j\Big)^{\alpha_j-1} \tag{A.12}$$

Now, by plugging the above equation into the workers equilibrium supply equation, i.e. (A.10) we obtain:

$$\mu_{kj} = \underbrace{m_k\chi_0(k)\Big(\frac{u_{kj}}{w_{k0}}\Big)^{\beta_j}(\gamma_{kj})^{\beta_j}}_{B(k)}(A_j)^{\beta_j}\Big(L_j\Big)^{\beta_j(\alpha_j-1)} \tag{A.13}$$

By multiplying both sides of (A.13) by γ_{kj}, and then sum over k we obtain the following equations:

$$L_j = \Big(L_j\Big)^{\beta_j(\alpha_j-1)} A_j^{\beta_j}\sum_{k=1}^K \gamma_{kj}B(k), \tag{A.14}$$

Then, by manipulating the above equations we obtain:

$$L_j = A_j^{\frac{\beta_j}{1+(1-\alpha_j)\beta_j}} \left\{ \sum_{k=1}^{K} \gamma_{kj} B(k) \right\}^{\frac{1}{1+(1-\alpha_j)\beta_j}}, \tag{A.15}$$

We can therefore plugging-in the above formulas in (A.12) and we obtain the following structural EWF:

$$\ln w_{kj} = \frac{1}{1+(1-\alpha_j)\beta_j} \ln A_j + \ln \gamma_{kj} - \frac{(1-\alpha_j)}{1+(1-\alpha_j)\beta_j} \ln \left\{ \sum_{k=1}^{K} (m_k \chi_0(k) (\gamma_{kj})^{1+\beta_j} \left(\frac{u_{kj}}{w_{k0}} \right)^{\beta_j} \right\}$$

Appendix B. Endogeneizing Amenities

In the main text we have considered that the workers have preferences for *exogenous amenities*, more precisely firms do not have power to set the level of amenities. In this section, we will consider a more general case where firms have power to set both the wages and also the level of amenities. Then amenities become now *endogenous*, while the costs of amenities firms are facing are *exogenous*. In such a context, the firm j optimization problem becomes:

$$\min_{w_{kj}, u_{kj}} \sum_{k \in \mathcal{K}} \big(w_{kj} + c_{kj} u_{kj} \big) \mu_{kj} \;\; s.t \; F^j(\mu_{\cdot j}) \geq Y^j, \quad w_{kj} \geq 0, \quad u_{kj} \geq 0$$

with

$$\mu_{kj} = m_k \frac{\partial G_{k\cdot}(v_{k\cdot})}{\partial v_{kj}}, \;\; (k,j) \in (\mathcal{K} \times \mathcal{J}); \tag{B.1}$$

where c_{kj} —such that $0 < c_{kj} \leq c < \infty$— denotes the cost of the amenities u_{kj}. The KKT first order necessary conditions lead to:

$$w_{kj} = \lambda_j F_k^j(\mu_{\cdot j}) \frac{\mathcal{E}_{kj}}{1 + \mathcal{E}_{kj} + \eta_{kj}} 1\{\lambda_j F_k^j > c_{kj} u_{kj}\}, \tag{B.2}$$

$$u_{kj} = \lambda_j F_k^j(\mu_{\cdot j}) \frac{\eta_{kj}/c_{kj}}{1 + \mathcal{E}_{kj} + \eta_{kj}} 1\{\lambda_j F_k^j > w_{kj}\}, \tag{B.3}$$

where $\eta_{kj} = \frac{u_{kj}}{s_{kj}} \frac{\partial s_{kj}}{\partial u_{kj}}$ is the elasticity of the type k labor supply relative to the amenities. Remark that when $\lambda_j F_k^j > \max(w_{kj}, c_{kj} u_{kj})$ we have only interior solutions. For now, let's focus only on the interior solutions. In the Logit economy and under Assumption 3, eqs (B.2) and (B.3) simplify to

$$w_{kj} = \psi_{kj} \frac{1 - s_{kj}}{3 - 2 s_{kj}}, \tag{B.4}$$

$$u_{kj} = \frac{\psi_{kj}}{c_{kj}} \frac{1 - s_{kj}}{3 - 2 s_{kj}}. \tag{B.5}$$

We obtain, therefore, the following relationship between the endogenous equilibrium wages and amenities:

$$\frac{w_{kj}}{u_{kj}} = c_{kj}. \tag{B.6}$$

We can still use our quasi-supply approach to obtain a tractable characterization of the equilibrium wages, amenities and matching distribution. Combining the Logit quasi-supply equation with eq (B.6) leads to: $s_{kj} = \frac{s_{k0}}{w_{k0}c_{kj}} w_{kj}^2$; and then, by plugging-in the latter into equation (B.4) and rearranging, we obtain the following cubic equation in w_{kj}:

$$w_{kj}^3 - \psi_{kj} w_{kj}^2 - 3\frac{w_{k0}c_{kj}}{s_{k0}} w_{kj} + \psi_{kj}\frac{w_{k0}c_{kj}}{s_{k0}} = 0. \tag{B.7}$$

It can been shown that this cubic equation has three real solutions, one non-positive and two non-negatives. The lowest non-negative solution can be shown to minimize firm j wage bill. This minimum non-negative wage provides the following EWF, equilibrium amenities function (EAF) and DMF:

$$w_{kj} = \frac{\psi_{kj}}{3} - 2\Big(\frac{w_{k0}c_{kj}}{s_{k0}} + (\frac{\psi_{kj}}{3})^2\Big)^{1/2} sin\left\{\frac{1}{3}\arcsin\left(\frac{\psi_{kj}/3}{\Big(\frac{w_{k0}c_{kj}}{s_{k0}} + (\frac{\psi_{kj}}{3})^2\Big)^{1/2}}\right)^3\right\}, \tag{B.8}$$

$$u_{kj} = \frac{\psi_{kj}}{3c_{kj}} - \frac{2}{c_{kj}}\Big(\frac{w_{k0}c_{kj}}{s_{k0}} + (\frac{\psi_{kj}}{3})^2\Big)^{1/2} sin\left\{\frac{1}{3}\arcsin\left(\frac{\psi_{kj}/3}{\Big(\frac{w_{k0}c_{kj}}{s_{k0}} + (\frac{\psi_{kj}}{3})^2\Big)^{1/2}}\right)^3\right\}, \tag{B.9}$$

$$s_{kj} = \frac{s_{k0}}{w_{k0}c_{kj}}\left[\frac{\psi_{kj}}{3} - 2\Big(\frac{w_{k0}c_{kj}}{s_{k0}} + (\frac{\psi_{kj}}{3})^2\Big)^{1/2} sin\left\{\frac{1}{3}\arcsin\left(\frac{\psi_{kj}/3}{\Big(\frac{w_{k0}c_{kj}}{s_{k0}} + (\frac{\psi_{kj}}{3})^2\Big)^{1/2}}\right)^3\right\}\right]^2, \tag{B.10}$$

for all $k \in \mathcal{C}^j, j \in \mathcal{J}$, and $\min(w_{kj}, u_{kj}) = s_{kj} = 0$ for all $k \in \overline{\mathcal{C}^j}, j \in \mathcal{J}$. Notice that, in the situation where $\lambda_j F_k^j < \max(w_{kj}, c_{kj}u_{kj})$, we have corner solutions with $w_{kj} = 0$ or $u_{kj} = 0$, in both cases $s_{kj} = 0$. Notice now, we have three endogenous equilibrium objects, the EWF, EAF & DMF, i.e. eq (B.8, B.9, B.10), but as before, they depend on the equilibrium object s_{k0}, which, for each type $k \in \mathcal{K}$ is the solution of eq (3.15).

As can be seen, our approach of using the quasi-supply instead of the supply to derive a tractable characterization of the equilibrium also apply to the case where the amenities are endogenously set. As before, this characterization of the equilibrium can allows us to easily analyze the existence and uniqueness of the our equilibrium matching model and also to derive informative closed form comparatives statistics. Let

$$\widetilde{H}(\widetilde{\varphi}_{k\cdot}, w_{k0}, s_{k0}) \equiv s_{k0} + \sum_{j\in\mathcal{J}_k} \frac{s_{k0}}{w_{k0}c_{kj}}\Bigg[\frac{\psi_{kj}}{3} - 2\Big(\frac{w_{k0}c_{kj}}{s_{k0}} + (\frac{\psi_{kj}}{3})^2\Big)^{1/2} sin\left\{\frac{1}{3}\arcsin\left(\frac{\psi_{kj}/3}{\Big(\frac{w_{k0}c_{kj}}{s_{k0}} + (\frac{\psi_{kj}}{3})^2\Big)^{1/2}}\right)^3\right\}\Bigg]^2 - 1$$

where $\widetilde{\varphi}_{k\cdot} = (\psi_{k\cdot}, c_{k\cdot})$ with $\psi_{k\cdot} \equiv (\psi_{k1}, ..., \psi_{kJ})$, $c_{k\cdot} \equiv (c_{k1}, ..., c_{kJ})$. We have the following result:

Lemma 2. *For all $k \in \mathcal{K}$ and $\varphi_{k\cdot} \in \big(\mathbb{R}_{>0}\big)^{2J}$ we have the following results:*

(i) $\lim_{s_{k0}\to 0} \widetilde{H}(\widetilde{\varphi}_{k\cdot}, w_{k0}, s_{k0}) = -1$, *and* $\widetilde{H}(\widetilde{\varphi}_{k\cdot}, w_{k0}, 1) > 0$.
(ii) $\frac{\partial \widetilde{H}(\widetilde{\varphi}_{k\cdot}, w_{k0}, s_{k0})}{\partial s_{k0}} > 0$.

Then, there exists a unique $s_{k0} \in (0,1)$ that solve the equation (3.15).

Lemma 2, shows that given a fixed set of parameters $\widetilde{\varphi} = (\widetilde{\varphi}_{1\cdot}, ..., \widetilde{\varphi}_{K\cdot})$ and w_{k0} there is a unique $s_{k0} \in (0,1)$ that solves equation (3.15) when both wage and amenities are endogenously determined by firms. This leads to the following result:

Theorem 3. *Consider a Logit economy where workers have preferences in the ARUM form and firms have linear productions functions (in other words Assumptions 3 and 4 hold), and Assumption 5 holds, then an equilibrium outcome of a many-to-one employee-employer matching model with imperfect competition where firms have power to set both wages and the level of amenities always exists and is unique. This equilibrium outcome is entirely characterized by the EWF, EAF & DMF, i.e., eqs (B.8, B.9, B.10) and the population constraints eq (2.11).*

Appendix C. Bootstrap CI for the ALC

In this section, we would like to give a brief description of the bootstrap percentile approach we use to construct the CI for the ALC. For doing so, we view $\phi_{\bar{k},j,\bar{k}',j'}$ as a function of $s_{\bar{k}j}$, $s_{\bar{k}'j'}$, $s_{\bar{k}'j}$, $s_{\bar{k}j'}$, $w_{\bar{k}j}$, $w_{\bar{k}'j'}$, $w_{\bar{k}'j}$, $w_{\bar{k}j'}$, $w_{\bar{k}j}$, and need to provide an accurate bootstrap version for them. In our model the number of firms is fixed so the asymptotic is on the workers side. Given that the ALC is constructed for each pair of firm j and j', the bootstrap version mimics this property by resampling among individuals within j and j'. Let $n(j,j')$ denotes the total number of workers in firm j and j', $n_{\bar{k}j}$ the number of type $\bar{k}$ workers in firm j, such that $\sum_{\bar{k}} n_{\bar{k}j} + \sum_{\bar{k}} n_{\bar{k}j'} = n(j,j')$. We have the following:

$$s_{\bar{k}j} = \frac{n_{\bar{k}j}}{m_{\bar{k}}} = \frac{n_{\bar{k}j}}{n(j,j')} \times \frac{n(j,j')}{M} \times \frac{M}{m_{\bar{k}}} = p_{\bar{k}j|j,j'} \times \frac{p(j,j')}{\mathfrak{m}_{\bar{k}}}$$

We propose as bootstrap analog for $s_{\bar{k}j}$ the following: $s^*_{\bar{k}j} = p^*_{\bar{k}j|j,j'} \times \frac{p(j,j')}{\mathfrak{m}_{\bar{k}}}$ where $p^*_{\bar{k}j|j,j'}$ is the bootstrap analog of $p_{\bar{k}j|j,j'}$ using resampling in the subpopulation of the total numbers of workers in firm j and j'. More precisely, $p^*_{\bar{k}j|j,j'} = \frac{n^*_{\bar{k}j}}{n(j,j')}$ where $n^*_{\bar{k}j}$ is the number of type $\bar{k}$ workers in firm j in the resampled subpopulation of size $n^*(j,j') = n(j,j')$. In each bootstrap resample, we denote by $w^*_{\bar{k}j}$ the average wage of type $\bar{k}$ workers in firm j. Using this process for each bootstrap we can compute:

$$\phi^{*,b}_{\bar{k},j,\bar{k}',j'} = \ln \frac{w^{*,b}_{\bar{k}j} w^{*,b}_{\bar{k}'j'}}{w^{*,b}_{\bar{k}j'} w^{*,b}_{\bar{k}'j}} + \ln \Big(\frac{2 - s^{*,b}_{\bar{k}j}}{1 - s^{*,b}_{\bar{k}j}}\Big)\Big(\frac{2 - s^{*,b}_{\bar{k}'j'}}{1 - s^{*,b}_{\bar{k}'j'}}\Big)\Big(\frac{1 - s^{*,b}_{\bar{k}'j}}{2 - s^{*,b}_{\bar{k}'j}}\Big)\Big(\frac{1 - s^{*,b}_{\bar{k}j'}}{2 - s^{*,b}_{\bar{k}j'}}\Big),\ b = 1, ..., B. \quad \text{(C.1)}$$

Therefore, the percentile bootstrap confidence interval for $\phi_{\bar{k},j,\bar{k}',j'}$ is given by $CI\big(\phi_{\bar{k},j,\bar{k}',j'}\big) = \Big(\phi^{*,\alpha/2}_{\bar{k},j,\bar{k}',j'}; \phi^{*,1-\alpha/2}_{\bar{k},j,\bar{k}',j'}\Big)$ where $\phi^{*,1-\alpha/2}_{\bar{k},j,\bar{k}',j'}$ denotes the $1-\alpha/2$ percentile of the bootstrapped distribution of $\phi^*_{\bar{k},j,\bar{k}',j'}$.

Appendix D. Proof of the main results

D.1. **Proof of Proposition 2.** To ease the evaluation of the Proof, let's recall here the *necessary and sufficient* conditions for local compatibility enumerated in Koning and Ridder (1994). Consider

J alternatives, with the following vector of stochastic utility $U_j = \tilde{v}_j + \epsilon_j,\;\; j = 1, ..., J$ and denotes $\tilde{v} = (\tilde{v}_1, ..., \tilde{v}_J)'$, $\iota_J = (1, ..., 1)'$. Also, let's denote by $\tilde{v}^j$ the $J-1$ sub-vector of $\tilde{v}^j$ that excludes the i^{th} component.

Necessary and sufficient conditions for local compatibility:

For all $\tilde{v} \in \tilde{\mathcal{V}}$ and $j = 1, ..., J$, $j \neq i = 1, ..., J$ we have:

(C-1) $\mathbb{P}_j(v) \geq 0$, and $\sum_{j=1}^{J} \mathbb{P}_j(v) = 1$,
(C-2) $\mathbb{P}_j(v) = \mathbb{P}_j(v + c\iota_j)$ for all $c \in \mathbb{R}$, (translation invariance)
(C-3) $\mathbb{P}_j(v)$ is differentiable respect to $\tilde{v}^j$,
(C-4) $\frac{\partial^k P_j}{\partial v^{jk}(v)} \geq 0$ with v^{jk} any k-subvector of v, (non-negativity)
(C-5) $\frac{\partial P_j}{\partial v_l}(v) = \frac{\partial P_l}{\partial v_j}(v)$. (symmetry)

According the only Theorem in Koning and Ridder (1994), (C-1)-(C-5) are the necessary and sufficient conditions for local compatibility. Now, consider our proposed probability choice:

$$s_{kj}(v_{k\cdot}) \;=\; \chi(k)\Big(e^{v_{kj}-v_{k0}} - a\Big)^{\beta_{kj}} 1\{j \in \mathcal{J}_k\} \tag{D.1}$$

$$s_{k0}(v_{k\cdot}) \;=\; 1 - \sum_{j \in \mathcal{J}_k} s_{kj}(v_{k\cdot}), \tag{D.2}$$

where $\chi_a(k) = \sum_{j\in\mathcal{J}_k}(b-a)^{\beta_{kj}}$ and $\widetilde{\mathcal{V}}_{k\cdot}(a,b) = \Big\{v_{k\cdot} : \ln a \leq v_{kj} - v_{k0} < \ln b, \quad j \in \mathcal{J}_k\Big\}$, where $0 \leq a < b < \infty$. For $j \notin \mathcal{J}_k$, (C-2) to (C-5) hold directly. Let focus on the case where $j \in \mathcal{J}_k$. We can easily see that (C-1), (C-2) and (C-3) hold by construction. For (C-4) remark that given our notation $v_j = -\tilde{v}_j$, so then we have $\frac{\partial s_{kj}(v_{k\cdot})}{\partial(-v_{k0})} = \frac{\partial s_{k0}(v_{k\cdot})}{\partial(-v_{kj})} = \beta_{kj}\chi(k)\Big(e^{v_{kj}-v_{k0}} - a\Big)^{\beta_{kj}-1} > 0$; and $\frac{\partial s_{kj}(v_{k\cdot})}{\partial(-v_{kl})} = 0$, for $j \neq l$, thus (C-4) and (C-5) holds. Also, it is worth-noting that strictly speaking the proof of the Theorem proposed by Koning and Ridder (1994) works for a connected regions formed by bounded intervals, which in our context applies when $a > 0$. In presence of arbitrary open sets, Koning and Ridder (1994)'s theorem does not directly holds, the compatibility has to be looking at cases by cases, as discussed Koning and Ridder (1994). In our context when a $\to 0$, $\lim_{a\to 0}\widetilde{\mathcal{V}}_{k\cdot}(a,b)$ is a connected region formed by right open intervals, and it can be showed that it works in our special case. The intuition is to follow Koning and Ridder (1994)'s approach to construct the joint distribution based on the observed choices that is compatible with ARUM for all $a > 0$. It can be seen that it remains a valid distribution when $a \to 0$.

References

[1] J. M. Abowd, F. Kramarz, and D. N. Margolis. High wage workers and high wage firms. *Econometrica*, 67:251–333, 1999.

[2] J. Alexander S. Kelso and V. P. Crawford. Job matching, coalition formation, and gross substitutes. *Econometrica*, 50(6):1483–1504, 1982.

[3] K. Arrow. *The Theory of Discrimination*. Princeton University Press, 1972.

[4] E. M. Azevedo and J. Hatfield. Complementarity and multidimensional heterogeneity in large matching markets. unpublished manuscript, 2013.

[5] G. S. Becker. *The Economics of Discrimination*. Economic Research Studies, 1957.

[6] D. Berger, K. Herkenhoff, and S. Mongey. Labor market power. Unpublished manuscript, 2019.

[7] S. Bonhomme, T. Lamadon, and E. Manresa. Discretizing unobserved heterogeneity. unpublished manuscript, 2017.

[8] S. Bonhomme, T. Lamadon, and E. Manresa. A distributional framework for matched employer employee data. Forthcoming at Econometrica, 2018.

[9] A. Börsch-Supan. On the compatibility of nested logit models with utility maximization. *Journal of Econometrics*, 43(3):373–388, 1990.

[10] C. Brown. Equalizing differences in the labor market. *The quarterly journal of economics*, 94(1):113–134, 1980.

[11] S. Caldwell and O. Danieli. Outside options in the labor market. Unpublished manuscript, 2018.

[12] D. Card, A. R. Cardoso, J. Heining, and P. Kline. Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics*, 36:S13–S70, January 2018.

[13] K. K. Charles and J. Guryan. Prejudice and wages: An empirical assessment of becker's the economics of discrimination. *Journal of Political Economy*, 116(51):773–809, 2008.

[14] E. Choo. Dynamic marriage matching: An empirical framework. *Econometrica*, 83(4):1373–1423, 2015.

[15] E. Choo and A. Siow. Who marries whom and why. *Journal of Political Economy*, 114(1):175–201, 2006.

[16] A. Dupuy and A. Galichon. A note on the estimation of job amenities and labor productivity. Unpublished manuscript, 2017.

[17] A. Dupuy, A. Galichon, S. Jaffe, and S. D. Kominers. Taxation in matching markets. Unpublished manuscript, 2018.

[18] F. Echenique, S. Lee, M. Shum, and M. B. Yenmez. The revealed preference theory of stable and extremal stable matchings. *Econometrica*, 81:153–171, 2013.

[19] J. Eeckhout and P. Kircher. Assortative matching with large firms. *Econometrica*, 86(1):85–132, 2018.

[20] J. Fox. Estimating matching games with transfers. *Quantitative Economics*, 9:1–38, 2018.

[21] D. Gale and H. Nikaido. The jacobian matrix and global univalence of mappings. *Mathematische Annalen*, 2(81–93), 1965.

[22] A. Galichon, S. Kominers, and S. Weber. Costly concessions: An empirical framework for matching with imperfectly transferable utility. Forthcoming at Journal of Political Economy, 2018.

[23] A. Galichon and B. Salanié. Cupid's invisible hand: Social surplus and identification in matching models. unpublished manuscript, 2015.

[24] B. Graham. Econometric methods for the analysis of assignment problems in the presence of complementarity and social spillovers. *Handbook of Social Economics*, 1B:965–1052, 2011.

[25] B. Graham. Comparative static and computational methods for an empirical one-to-one transferable utility matching model. *Advances in Econometrics*, 31:153–181, 2013.

[26] J. W. Hatfield and P. R. Milgrom. Matching with contracts. *American Economic Review*, 95(4):913–935, 2005.

[27] J. J. Heckman and B. E. Honoré. The empirical content of the roy model. *Econometrica*, 58(5):1121–1149, 1990.

[28] R. H.Koning and G. Ridder. On the compatibility of nested logit models with utility maximization: A comment. *Journal of Econometrics*, 63(2):389–396, 1994.

[29] R. H. Koning and G. Ridder. Discrete choice and stochastic utility maximization. *Econometrics Journal*, 6:1–27, 2003.

[30] T. Lamadon, M. Mogstad, and B. Setzler. Imperfect competition and rent sharing in the u.s. labor market. unpublished manuscript, 2017.

[31] A. Manning. Imperfect competition in the labor market. *Handbook of Labor Economics,*, 4:973–1041.

[32] A. Manning. *Monopsony in motion: Imperfect Competition in Labor Markets*. Princeton University Press, 2003.

[33] D. McFadden. Conditional logit analysis of qualitative choice behavior. In *Frontiers in Econometrics*, pages 105–142. Academic Press, New York, zarembka edition, 1993.

[34] D. L. McFadden. *Economic Models of Probabilistic Choice*, chapter Structural Analysis of Discrete Data with Econometric Applications. MIT Press, 1981.

[35] D. L. McFadden. *Modeling the Choice of Residential Location*, chapter Spatial Interaction Theory and Residential Location. North-Holland, 1987.

[36] I. Mourifié. A marriage matching function with flexible spillover and substitution patterns. *Economic Theory*, pages 1–41, 2018.

[37] I. Mourifié, M. Henry, and R. Meango. Sharp bounds and testability of a roy model of stem major choices. unpublished manuscript, 2018.

[38] I. Mourifié and A. Siow. The cobb douglas marriage matching function: Marriage matching with peer effects. unpublished manuscript, 2017.

[39] T. Parthasarathy. *On Global Univalence Theorems*. Springer, March 1983.

[40] F. Postel-Vinay and J.-M. Robin. Equilibrium wage dispersion with worker and employer heterogeneity. *Econometrica*, 70(6):2295–2350, 2002.

[41] M. Pycia. Stability and preference alignment in matching and coalition formation. *Econometrica*, 80(1):323–362, 2012.

[42] M. R. Ransom and V. E. Lambson. Monopsony, mobility, and sex differences in pay: Missouri school teachers. *The American Economic Review, papers and proceedings*, 101(3):454–0459, 2011.

[43] M. R. Ransom and R. L. Oaxaca. New market power models and sex differences in pay. *Journal of Labor Economics*, 28(2):267–289, 2010.

[44] J. Robinson. *The Economics of Imperfect Competition.* Palgrave Macmillan, 1933.

[45] S. Rosen. Hedonic prices and implicit markets: Product differentiation in pure competition hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82(1):34–55, 1974.

[46] S. Rosen. The theory of equalizing differences,. *Handbook of Labor Economics,*, 1:641–692, 1986.

[47] A. Roy. Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3:135–146.

[48] I. Sorkin. Ranking firms using revealed preference. *The quarterly journal of economics*, 133(3):1331–1393, 2018.

[49] C. Taber and R. Vejlin. Estimation of a roy/search/compensating differential model of the labor market. unpublished manuscript, 2016.

[50] A. Tversky. Elimination by aspects: A theory of choice. *Psychological Review*, 79(4):281–299, 1972.

[51] M. S. Xiaoxia Shi and W. Song. Estimating semi-parametric panel multinomial choice models using cyclic monotonicity. *Econometrica*, 86(2):737–761.