How Monopsonistic and Monopolistic Competition Affects Wage Disparities?

Jacques-François Thisse* Eric Toulemonde†

3rd April 2015

Abstract

We develop a general equilibrium model in which firms employ different types of labor. Firms are endowed with market power that allows them to be price-makers and wage-setters. First, firms face upward sloping labor supplies because idiosyncratic non-pecuniary conditions interact with wages in workers' decisions to work for specific firms. Second, we pin down the existence of a double exploitation of labor whose intensity depends on the interaction between the product and labor markets. Third, the heterogeneity within each type of labor implies that the high-productive workers tend to be overpaid, whereas the low-productive workers would be underpaid. However, intensifying competition on the goods market shrinks the discrepancy between wages and workers' productivity. Last, we offer a theory of differential discrimination in which gender pay inequality varies with women's family status.

Keywords: worker heterogeneity; monopsonistic competition; monopolistic competition; wage dispersion

JEL Classification: D33, J31, J42, J71, L13

*CORE-UCLouvain (Belgium), Higher School of Economics (Russia) and CEPR.
†Université de Namur (Belgium), CORE and IZA.

1 Introduction

A large literature has documented the following facts: (i) the wage gap between the skilled and the unskilled has grown rapidly over the past 30 years, (ii) women are discriminated against on labor markets, (iii) gender discrimination shrinks as the product market becomes more competitive, and (iv) wage discrimination within the same group of workers occurs according to their familial or marital status. A variety of approaches have been developed to rationalize those facts, often within partial equilibrium frameworks. As a consequence, the resulting explanations are often scattered and lack unity. The main contribution of this paper is to provide a broader setup that knits within a unifying and tractable framework the various approaches used in the literature.

We develop a general equilibrium model in which firms are endowed with market power on the product and labor markets. The central tenet of this paper is that blending imperfections on both the labor and product markets yields new and insightful results about the distribution of earnings across different types of production factors. In perfectly competitive markets, the ratio of wages to productivity of each production factor is equal to one. Imperfect competition on the product market alone lowers the value of the ratio but does not affect the equalization of the ratio across production factors. Imperfect competition on the labor market not only changes the ratio but it also affects it differently across factors, even under free entry. Our model shows that if the ratio decreases for one factor (e.g., labor, female workers, unskilled workers), it must increase for another factor (e.g., capital, male workers, skilled workers).

Although a comprehensive general equilibrium model with strategic interactions on these two markets has so far been out of reach, it is possible to gain relevant insights by considering a large number of firms (formally, a continuum). Under these circumstances, a firm has market power but is negligible to the market. Regarding the product market, we consider two different settings that allow us to capture different market environments. In the first one, the number of firms is exogenous. As a consequence, market power generates a rent accruing to firms. In the second setting, the rent is eliminated by free entry on the product market.

We model the product market using a model in which (i) the markup is variable, (ii) preferences may differ across types of labor, and (iii) substitution between different types of labor is allowed. As for the labor market, we build on the growing evidence that firms face elastic labor supplies (Ashenfelter et al., 2010). To be precise, we assume that idiosyncratic non-pecuniary conditions interact with wages in workers' decisions to accept offers made by specific firms. In other words, different workers view jobs offered by different firms as bundles of hedonic attributes that provide them with more or less satisfaction. Since workers make mutually exclusive and indivisible job choices, discrete choice theory provides us with an appropriate tool to model the actual matching value between a worker and a firm. Specifically, we assume that heterogeneity differs between types of labor and is captured by the logit model within each type.

Our main findings may be summarized as follows. Starting with an exogenous number of firms, we show that *a firm's labor supply curve is elastic when workers of a given type are heterogeneous.* Each worker having a most-preferred employer, firms may set a wage lower than workers' marginal product value while attracting their captive labor pool. Workers are paid below their marginal value product for a second reason. Since firms are price-makers on the product market, they evaluate workers' marginal productivity at the marginal revenue, which is lower than the market price. Hence, once it is recognized that both the product and labor markets operate under imperfect competition, the market delivers an outcome that involves a monetary transfer from the workers to the firms. It is worth stressing here that the intensity of this double exploitation of labor depends on the interaction between the product and labor markets. For example, the wage markdown rises when workers' preferences for hedonic job attributes grows relatively to that for consumption goods. However, a price drop on the product market weakens this effect by reducing firms' monopsonistic power on the labor market.

Second, *the wage gap exceeds the productivity gap between any two different types of labor.* Indeed, a worker's hedonic wage is equal to her actual wage plus the pecuniary evaluation of her job's attributes. As a consequence, the market works as if the more productive employees were more sensitive to wage differences than the less productive employees, even when they are equally heterogeneous. In other words, heterogeneity within each type of labor magnifies productivity differences between types. Note, however, that the premium paid to the more productive workers decreases when the product market becomes more competitive. In this sense, a more competitive product market lessens the divergence between wages and productivity.

Third, even when workers of different types have the same productivity, wage dispersion may still arise because those workers need not have the same degree of heterogeneity in their preferences for jobs. For instance, in Denmark, Eriksson and Kristensen (2014) find that women value flexibility significantly more than men, while in Spain de la Rica et al. (2010) find that monopsonistic features, which could be related to women' lower labor mobility due to housework, explain the gender wage gap. We also rationalize the fact that gender wage inequality is not uniform and varies between women with children and women without children (Polachek, 2014). We thus offer a theory of *differential discrimination*. However, we will show that these results need qualification when preferences for goods strongly differ across types of labor. Thus, ignoring difference in preferences for the final goods may lead to inaccurate conclusions in empirical studies of labor markets.

Fourth, when the number of firms is determined by free entry, income transfers from workers to firms vanish because profits are zero. Contrary to general belief, this does not wash out discrimination among workers' types. Rather, monopsony power takes the concrete form of implicit transfers across different types of labor. Because the above-mentioned magnification of productivity differences also holds under free entry, this effect expresses itself through an income transfer from the low-productive workers to the high-productive workers. To put it differently, *the high-skilled workers are overpaid, whereas the low-skilled workers are underpaid*. What is more, since the skilled tend to display a growing geographical mobility relative to the unskilled (Docquier and Rapoport, 2012; Moretti, 2012), workers' differential mobility is likely to be part of the explanation for the growing wage inequality between these two types of workers. More generally, we show that workers who value less jobs' hedonic attributes gain at the expense of those for whom these attributes are more important. In the limit, even when they share identical observable characteristics in every other aspect, a group of workers that value more such attributes than another will be discriminated against.

Note, finally, that our model is versatile enough to shed light on two important issues related to labor economics, namely the existence of an urban wage premium and the declining share of labor in the gross domestic product of developed countries.

Related literature. Kim (1989), Bhaskar and To (1999, 2003), Marimon and Zilibotti (1999), and Hamilton et al. (2000) build on Salop (1979) to model heterogeneous firms competing to attract heterogeneous workers. The introduction of strategic considerations in the labor market renders the analysis much more involved, thus leading these authors to consider the product side as perfectly competitive. The majority of models that blend imperfections on the product and labor markets use a setting in which wages are bargained between workers and employers (Blanchard and Kiyotaki, 1987; Blanchard and Giavazzi, 2003). In Helpman et al. (2010), the labor market is characterized by search and matching frictions, which are modeled following the standard Diamond-Mortensen-Pissarides approach. However, unlike us, those various authors consider a single type of labor and, therefore, cannot analyze the distributional consequences of imperfect markets between types of labor, such as skill, gender and marital status.

The paper proceeds as follows. We present the model in Section 2. In Section 3, we characterize the equilibrium when the number of firms is exogenous, while the subsequent section considers the case of free entry. The last section summarizes our main policy implications and discusses possible extensions.

2 The model

We consider an economy endowed with one sector and Θ types of labor, such as the skill level of workers, their age category, gender and ethnicity. There are L_θ workers of type $\theta = 1, ...\Theta$ and each worker has one unit of her type of labor. The total population is denoted by $L = \Sigma_\theta L_\theta$, while $\mathbf{L}$ denotes the vector $(L_1, ..., L_\Theta)$. There are two goods. The homogeneous good is unproduced (land) and its supply H is perfectly inelastic; it is used as the numéraire. Each worker is endowed with H/L units of this good. The differentiated good is produced and made available as a continuum of varieties of mass N; each variety is denoted by $i \in [0, N]$.

2.1 Workers

Workers are heterogeneous in their preferences for consumption goods and for jobs.

(i) **Consumption**. Workers of type $\theta = 1, ..., \Theta$ share the same strictly quasi-concave utility function:

$$U_\theta \left(h, u \left(x \left(i\right); i \in [0, N]\right)\right), \tag{1}$$

where h is the consumption of the numéraire and $x(i)$ the consumption of variety i. The subutility u is strictly quasi-concave and symmetric in the set of varieties, while the utility U_θ is θ-specific. In other words, workers endowed with different types of labor may have different attitudes toward consumption. For analytical simplicity, we assume that workers endowed with the same type of labor have the same preferences for goods.

In what follows, we will focus on symmetric equilibria. For this to happen, consumers of the same type must have the same income, which holds when profits are uniformly distributed across consumers. A θ-worker hired by firm i earns a nominal wage $w_\theta \left(i\right)$ and has a budget constraint given by

$$\int_0^N p\left(j\right) x\left(j\right) \mathrm{d}j + h = I_\theta \left(i\right) \equiv w_\theta \left(i\right) + \frac{H}{L} + \frac{1}{L}\int_0^N \pi\left(j\right) \mathrm{d}j,$$

where $\pi\left(j\right)$ is the profit made by firm j (see below for more details).

Individual preferences are such that the θ-workers aggregate demand for variety j, denoted $x_\theta\left(j\right)$, decreases with $p\left(j\right)$ and increases with the sum of the θ-workers' incomes:

$$I_\theta \equiv \int_0^N I_\theta \left(i\right) \mathrm{d}i.$$

Observe that the income I_θ depends on the wage w_θ that depends on all workers' expenditure, which in turn varies with the income I_θ of the θ-type of workers. Our model is thus a full-fledged general equilibrium model.

The total expenditure on the differentiated good is given by

$$E \equiv \int_0^N \left[\sum_\theta p\left(j\right) x_\theta \left(j\right)\right] \mathrm{d}j.$$

It will be convenient to illustrate our results in the widely used case of a Cobb-Douglas utility nesting a CES subutility:

$$\begin{aligned} U_\theta(h, M) &= \frac{M^{\mu_\theta} h^{1-\mu_\theta}}{(\mu_\theta)^{\mu_\theta} (1-\mu_\theta)^{1-\mu_\theta}} & 0 < \mu_\theta < 1 \\ M &\equiv \left[\int_0^N (x(i))^{\frac{\sigma-1}{\sigma}} \mathrm{d}i \right]^{\frac{\sigma}{\sigma-1}} & \sigma > 1. \end{aligned} \tag{2}$$

In this case, which we call **C-D-C**, the total expenditure on the differentiated good is given by

$$E = \sum_\theta \mu_\theta I_\theta,$$

while the aggregate demand for variety i is given by

$$q(i) = \left[\frac{p(i)}{P} \right]^{-\sigma} \frac{E}{P},$$

where

$$P \equiv \left[\int_0^N \left(p(j)^{-(\sigma-1)} \right) \mathrm{d}j \right]^{\frac{-1}{\sigma-1}}$$

is the CES price index of the differentiated good. In what follows, $\mathbf{P}_\theta \equiv P^{\mu_\theta} \cdot 1^{1-\mu_\theta}$ denotes the *general* price index faced by the θ-workers, which varies across types of labor because workers differ by the earning share they spend on the differentiated good.

(ii) **Jobs.** Workers sharing the same type are heterogeneous in their perception of the hedonic attributes associated with a particular firm/job or, equivalently, the importance of these attributes relative to their wages is worker-specific. Formally, this is modeled by assuming that the random indirect utility of a θ-worker employed in firm i is given by

$$\tilde{V}_\theta(i) \equiv V_\theta(i) + \varepsilon_\theta(i),$$

where $V_\theta(i)$ denotes the indirect utility the worker enjoys from consuming the homogeneous and differentiated goods and earning the wage $w_\theta(i)$ in firm i. In the **C-D-C** case, $V_\theta(i)$ boils down to

$V_\theta(i) = I_\theta(i) / \mathbf{P}_\theta$.

A workers' idiosyncratic taste for a job provided by firm i is given by the realization of the zero-mean random variable $\varepsilon_\theta(i)$, which is known to the worker but unobservable by the firms. A worker chooses the firm that grants her with the highest random indirect utility $\tilde{V}_\theta(i)$ given by

$$\max_i \left[V_\theta(i) + \varepsilon_\theta(i)\right],$$

which depends on the wages set by firms and their hedonic attributes. This implies that, even when the θ-workers who choose to work in firm i are paid the same wage, they value differently the hedonic attributes of this firm and thus enjoy different welfare levels.

In this paper, we follow the discrete choice models of market competition and assume that the random variables $\varepsilon_\theta(i)$ are independently and identically distributed according to the Gumbel distribution. This implies that the probability she chooses to work in firm i is given by the *continuous logit* (Ben-Akiva et al., 1985; Dagsvik, 2002):

$$\mathbb{P}_\theta(i) = \frac{\exp \frac{V_\theta(i)}{\gamma_\theta}}{\int_0^N \exp \frac{V_\theta(j)}{\gamma_\theta} \mathrm{d}j}, \tag{3}$$

where γ_θ stands for the standard-deviation of $\varepsilon_\theta(i)$ (up to the numerical factor $\pi/\sqrt{6}$). In (3), γ_θ is an index that captures the heterogeneity of workers who react differently to the same wage schedule within the θ-type of labor. Alternatively, γ_θ may be interpreted as an inverse measure of θ-workers' *inter-firm and/or geographical mobility*: a larger γ_θ implies that a smaller share of θ-workers is willing to change jobs in response to a wage cut. Throughout the remaining of the paper, we assume that γ_θ is small enough for all the expressions derived below to be positive.

2.2 Firms

The differentiated good is produced under increasing returns and monopolistic competition. Each firm supplies a single variety and each variety is produced by a single firm. Consequently, a variety may be identified by its producer $i \in [0, N]$. Firm i hires $\ell_\theta(i) \geq 0$ workers of type $\theta = 1, ...\Theta$, and this firm's production function is given by a linear homogeneous function $F[\mathbf{l}(i)]$, where $\mathbf{l}(i) \equiv (\ell_1(i), ..., \ell_\Theta(i))$. The output of firm i is split between the fixed requirement f needed to undertake production and the quantity $q(i)$ offered to consumers:

$$q(i) + f = F[\mathbf{l}(i)]. \tag{4}$$

Each firm i chooses the wage $w_\theta(i)$ it pays to each type of labor θ and attracts

$$\ell_\theta(i) = L_\theta \mathbb{P}_\theta(i) \tag{5}$$

θ-workers. Because each firm is negligible to the market, when choosing the salary $w_\theta(i)$ it will pay to the θ-workers, firm i treats accurately the denominator of (3) as a given, very much like firms view the price index of the product market as a parameter in the Dixit-Stiglitz model. In contrast, the numerator of (3) is affected by the choice of $w_\theta(i)$. Therefore, firms face *monopsonistic competition* on the labor market.

By choosing the wage $w_\theta(i)$ firm i determines its employment level $\ell_\theta(i) = L_\theta \mathbb{P}_\theta(i)$, which pins down the firm's output $q(i) = F[\mathbf{l}(i)] - f$. Through its inverse demand function, this in turn determines the price $p(i)$ at which firm i sells its variety. Hence, though firms operate on the product market as if there were monopolistic competition on this market, they cannot choose their prices $p(i)$ independently of their wages $w_\theta(i)$. However, the price at which firm i can sell its variety is endogenous and determined by the demand for its variety. As a consequence, the equilibrium wage is determined by the competitive conditions on the labor and product markets through the demands for varieties.

Firm i maximizes its profits given by

$$\pi(i) = p(i)\, q(i) - \sum_\theta w_\theta(i)\, \ell_\theta(i), \tag{6}$$

subject to the production function (4) and the inverse demand function obtained from (1). Formally, firm i's profit-maximizing wages solve the following profit-maximizing conditions:

$$\frac{d\pi(i)}{dw_\theta(i)} = p(i)\left[1+\frac{\partial p(i)}{\partial q(i)}\frac{q(i)}{p(i)}\right]\frac{\partial q(i)}{\partial \ell_\theta(i)}\frac{\partial \ell_\theta(i)}{\partial w_\theta(i)} - \ell_\theta(i)\left[1+\frac{\partial \ell_\theta(i)}{\partial w_\theta(i)}\frac{w_\theta(i)}{\ell_\theta(i)}\right] = 0.$$

This expression says that the equilibrium wage $w_\theta(i)$ set by firm i is such that the additional revenue earned by hiring θ-workers at a higher wage is equal to the increase in the wage bill borne by the firm. Note that the corresponding increase in marginal cost stems from the heterogeneity of the θ-workers while the marginal revenue differs from the market price because of firm i's market power on the product market.

3 Equilibrium under a given number of firms

3.1 Wage equation

Since firm i is negligible to the labor market, it accurately treats the denominator of (3) parametrically. Hence,

$$\frac{\mathrm{d}\ell_\theta(i)}{\mathrm{d}w_\theta(i)} = \frac{V'_\theta(i)}{\gamma_\theta}\ell_\theta(i), \tag{7}$$

where $V'_\theta(i)$ denotes the marginal indirect utility of a θ-type worker employed by firm i. The elasticity of firm i's labor supply is thus given by

$$e_\theta(i) = \frac{V'_\theta(i)}{\gamma_\theta}w_\theta(i).$$

Thus, although the market supply of labor is perfectly inelastic, each firm faces a supply curve with a finite elasticity because the θ-workers are heterogeneous ($\gamma_\theta > 0$). Everything else being equal, *the more heterogeneous the θ-workers, the smaller the elasticity of the supply curve.* By contrast, the labor supply curve is infinitely elastic when θ-workers care only about their wage ($\gamma_\theta = 0$).

In the special case where preferences are given by (2), the elasticity of firm i's labor supply is

$$e_\theta(i) = \frac{w_\theta\left(i\right)}{\gamma_\theta \mathbf{P}_\theta},$$

which implies that the elasticity rises when the price index falls, e.g. through a higher elasticity of substitution σ. The intuition is easy to grasp. In choosing their employer, workers face a trade-off between the hedonic job attributes (the value of $\varepsilon_\theta\left(i\right)$) and the nominal wage. Since the indirect utility rises when prices are lower, workers put more weight on their nominal wage than on the hedonic attributes. This increases the elasticity of the labor supply schedules and weakens firms' monopsony power. Observe that the same holds when the θ-workers become more homogeneous. In the limit, when they are homogeneous ($\gamma_\theta = 0$), we fall back on the standard case of a perfectly elastic labor supply. This highlights the role of workers' heterogeneity for firms to face elastic labor supplies.

The following proposition is a summary.

Proposition 1. *A firm's labor supply gets more elastic as the product market becomes more competitive, the workers more homogeneous, or both.*

We now come to the wage determination. Using (7), we obtain the following wage equation:

$$w_\theta^*\left(i\right) = p\left(i\right)\left[1 + \frac{q(i)}{p\left(i\right)}\frac{\partial p\left(i\right)}{\partial q\left(i\right)}\right] F_\theta'(i) - \frac{\gamma_\theta}{V_\theta'(i)}, \tag{8}$$

where $F_\theta'(i)$ denotes the derivative of F (and thus of $q\left(i\right)$) with respect to $\ell_\theta\left(i\right)$. Note that $V_\theta'(i)$ depends wages and market prices, so that the wage equation is *not* additively separable in the effects of firms' market power and workers' heterogeneity.

The equilibrium wage of θ-workers differs from their marginal value product, $p(i)F_\theta'(i)$. To be precise, the exploitation of workers has *two* sources. First, firms use their monopoly power on the product market to set a markup equal to $1/(1 + (q/p)(\partial p/\partial q)) > 1$ times the marginal cost. As a consequence, the θ-workers' wage is evaluated at to the marginal value product times the inverse markup, that is, *wages are determined by the marginal revenue*, not by the market price. Econometric estimations undertaken by Anderson and van Wincoop (2004) show that the elasticity $-1/(q/p)(\partial p/\partial q)$ varies from 5 to 10, which suggests a first exploitation rate of about 10 to 20 percent of the marginal value product.

Second, since θ-workers are heterogeneous in their preferences for employers, firms exercise their clout to pay the θ-workers a wage smaller than the marginal revenue. The wage drop is equal to $\gamma_\theta/V_\theta'(i)$, and thus depends on market prices. It follows from (3) that *a higher marginal indirect utility incentivizes workers to seek better paid jobs rather than better hedonic attributes.* Indeed, since lower prices increase the value of the marginal indirect utility, tougher competition on the product market reduces the negative effect that the heterogeneity of preferences for jobs exercises on wages.

As noticed by Boal and Ransom (1997), Pigou (1924) used the ratio $1/e_\theta = (\gamma_\theta/V_\theta')/w_\theta$ to measure labor exploitation. According to recent estimations the firm's labor supply elasticity would range from 2 to 4 (Manning, 2003; Ashenfelter et al., 2010), suggesting a second exploitation rate varying from 25 to 50 percent of the observed wage. Accordingly, the degree of exploitation of labor is far from being negligible. There is 'double exploitation' of labor, that is, an income transfer away from workers to firms, generated by the interaction between the product and labor markets. It is worth stressing that the double exploitation is here the unintentional consequence of decisions made by a great number of firms and workers.

Thus far, the equilibrium price $p(i)$ is undetermined. Given the symmetry of our setting, we find it natural to focus on symmetric market outcomes, which helps us to find how the price is determined. Under symmetry,

$$\ell_\theta(i) = L_\theta/N \qquad q(i) = q \qquad p(i) = p.$$

Using (4) and $\ell_\theta(i) = L_\theta/N$, the equilibrium output

$$q^* \equiv F(L_1/N, ..., L_\Theta/N) - f$$

decreases with the number of firms.

As for the equilibrium price, it must be such that the value of production, $p\left(F\left(\mathbf{L}\right)-Nf\right)$, is equal to the total expenditure on the differentiated varieties, E:

$$p^{*}=\frac{E}{F\left(\mathbf{L}\right)-Nf}. \tag{9}$$

The equilibrium price is linear in workers' total expenditure on the differentiated product, which varies with the incomes of all types of labor. The total income of θ-workers is equal to the sum of their wages, share of total profits and initial endowments of the homogeneous good:

$$I_{\theta}=L_{\theta}w_{\theta}+\frac{L_{\theta}}{L}\left(E-\sum_{t}w_{t}L_{t}\right)+\frac{L_{\theta}}{L}H.$$

Since I_{θ} is endogenous, the equilibrium price and wages are implicit functions of N. Yet, although an explicit solution for w_{θ}^{*} seems to be out of reach, the above expressions will be used below to better understand the interactions between the product and labor markets.

3.2 The equilibrium in the C-D-C case

The above expressions become easier to interpret in the **C-D-C** case, assuming further that workers are homogeneous in their taste for goods, $\mu_{\theta}=\mu$. In this case, the equilibrium price is given by

$$p^{*}=\frac{\mu}{1-\mu}\frac{H}{F\left(\mathbf{L}\right)-Nf}, \tag{10}$$

which is independent of σ and γ_{θ}. This expression shows that *a larger number of firms yields a higher market price*, a result that runs against the conventional wisdom, which states that entry leads to lower market prices. This may be explained as follows. A larger number of firms makes competition tougher on the product market, thus pushing the market price downward. However, when the labor force L remains constant, the entry of new firms reduces the employment and output of each firm $\left(F\left(\mathbf{L}/N\right)-f\right)$, which in turn fosters a higher price on the product market.

What (10) shows is that the latter effect dominates the former when workers share the same preferences for goods. There is no reason to expect this result not to hold when workers have heterogeneous tastes for goods.

Wages are now explicitly given by

$$w_\theta^* = \frac{\sigma - 1}{\sigma} p^* F_\theta' - \gamma_\theta \mathbf{P}^*, \quad \text{where} \quad \mathbf{P}^* = N^{-\mu/(\sigma-1)} (p^*)^\mu. \tag{11}$$

This expression shows that the equilibrium wage of θ-workers depends on the product market for the following three reasons. First, wages depend on the market price p^*, which determines the marginal value product of labor ($p^* F_\theta'$).

Second, $(\sigma - 1)/\sigma < 1$ represents firms' relative *markdown* generated by monopolistic competition on the product market. As observed by Robinson (1933), when there is imperfect competition on the product market, the equilibrium wage is smaller than the competitive wage, the markdown being given here by $(\sigma - 1)/\sigma < 1$. Even in the absence of imperfections on the labor market ($\gamma_\theta = 0$), imperfect competition on the product market translates into a wage smaller than the competitive wage because firms strives to produce less and, accordingly, hire fewer workers. Moreover, as the degree of firms' monopoly power on the product market rises, that is, σ falls, it follows from (10) and (11) that the equilibrium wage decreases. The argument is straightforward. Since firms further reduce their output, they hire fewer workers, thereby making competition on the labor market softer.

Third, θ-workers' hedonic job attributes are measured by $\gamma_\theta \mathbf{P}^*$, which depends on the price of the differentiated good and the income share μ spent on this good. This second markdown increases with the market price p^*. Indeed, a higher price reduces the marginal indirect utility of the differentiated good, and thus increases the relative value of the hedonic attributes. As a consequence, workers are willing to trade these attributes for a lower wage.

Consider now a positive shock on the market price p^*, such as a higher degree of product differentiation. This gives rise to the following two opposite effects. On the one hand, a higher market price has a direct positive impact on wage - see the first term of w_θ^*. On the other hand, this increases the price index, which raises the degree of labor exploitation - see the second term of w_θ^*. Differentiating (11) with respect to p^* shows that the former effect is stronger than the latter one: the equilibrium wages rise with the market price. In addition, an increase in f being equivalent to an hike in p^*, workers earn a higher wage in industries where the degree of increasing returns is higher. Indeed, when f rises, the quantity produced for consumption gets smaller. This in turn allows firms to sell at a higher price, and thus to pay a higher wage.

Note, finally, that w_θ^* as given by (11) is a fairly involved function of N. Nevertheless, we show in Appendix that *the equilibrium wage always increases with the number of firms*. Indeed, a larger number of firms operating under increasing returns makes competition tougher on the labor market, and thus leads to a higher wage for any given σ.

In sum, the above discussion shows that the interaction between the product and labor markets is fairly involved, even in the **C-D-C** case.

3.3 The wage structure

Consider now any two different types of labor, $\theta = k, l$. Then, we have:

$$w_k^* = \frac{F_k'}{F_l'}\left(w_l^* + \frac{\gamma_l}{V_l'}\right) - \frac{\gamma_k}{V_k'}. \tag{12}$$

When all workers are homogeneous ($\gamma_k = \gamma_l = 0$), we fall back on the well-known equality between the wage ratio w_k^*/w_l^* and the marginal productivity ratio F_k'/F_l'. By contrast, when workers are heterogeneous ($\gamma_k > 0$ and $\gamma_l > 0$), this equality ceases to hold. The relationship between the wage ratio and the marginal productivity ratio now depends on γ_θ and V_θ', which reflect the different attitudes of workers toward job attributes (γ_θ) and consumption goods (V_θ'). Hence, workers' heterogeneity suffices to break down the classical relationship between the productivity and wage ratios. In other words, workers' heterogeneity generates *direct interactions* between the labor and product markets through the parameters γ_θ/V_θ', which has redistributional implications between types of labor. For example, the heterogeneity of type-l workers positively affects type-k

workers when these ones are homogeneous ($\gamma_k = 0$ and $\gamma_l > 0$):

$$\frac{w_k^*}{F_k'} = \frac{w_l^*}{F_l'} + \frac{\gamma_l}{V_l'} > \frac{w_l^*}{F_l'}.$$

To further illustrate, we first consider the **C-D-C** setting in which workers share the same preferences for goods and for jobs ($\mu_k = \mu_l$ and $\gamma_k = \gamma_l = \gamma$), but differ in productivity ($F_k' > F_l'$). In this case, (12) becomes:

$$w_k^* = \frac{F_k'}{F_l'} w_l^* + \left(\frac{F_k'}{F_l'} - 1\right) \gamma \mathbf{P}^*.$$

Since $F_k' > F_l'$, this expression implies the following result.

Proposition 2. *More workers' heterogeneity, a greater monopoly power on the product market, or both exacerbate the wage difference between the more and less productive types of labor.*

Intuitively, the relative value of hedonic job attributes is lower for the high-wage workers than for the low-wage workers, thus making the high-wage workers more sensitive to wage differences than the low-wage workers. Moreover, the premium paid to the more productive workers rises when competition on the product market is relaxed, perhaps through product differentiation. To put it differently, intensifying competition on the goods market shrinks the discrepancy between wages and workers' productivity.

This concurs with MacDonald and Reynolds (1994) who found substantial evidence that the wedge between the wage and the marginal value product is higher for a young baseball player than for an experienced player. These authors showed that salary differences between first and second rank performers greatly exaggerate talent differences. It is worth stressing that the productivity level of workers reflects here their skill level as well as their relative scarcity. In the same spirit, since academics are likely to display fairly heterogeneous preferences in taste for jobs, *universities will pay disproportionately high salaries to the super-stars, while underpaying the others.*

We now consider the reverse case in which all workers have the same productivity ($F_k' = F_l'$) but different attitudes toward non-monetary job attributes ($\gamma_k < \gamma_l$). It is then readily verified that (12) is equivalent to

$$w_k^* - w_l^* = (\gamma_l - \gamma_k)\mathbf{P}^* > 0. \tag{13}$$

Ever since Gary Becker, there is a wide consensus in economic theory that increasing competition tends to reduce wage discrimination. And indeed, (13) shows that wage discrimination between the two types of labor declines as the product market gets more competitive, that is, when the price index falls. For example, Black and Strahan (2001) find that the gap between men's and women's wages shrunk after the deregulation of the banking sector in the US, but did not disappear. Indeed, the wage gap never vanishes as long as γ_k and γ_l differ.

Since the expression (13) remains valid under free entry, that is, when firms' profits are zero, we may conclude that wage discrimination is not caused by the sole existence of a rent on the product market allocated by firms among different groups of workers. Rather, *heterogeneity is the cause and discrimination the consequence.* To be precise, wage discrimination reflects the heterogeneity of workers between groups of workers, while discrimination is exacerbated when profits are positive and/or the market becomes less competitive.

4 The free-entry equilibrium

In this section, we assume that firms are free to enter and exit the market and study how this process affects the product and labor markets. Substituting (8) and (9) into (6), we obtain the zero-profit condition:

$$N\pi(i) = E - \sum_{\theta} w_\theta L_\theta = -\frac{E}{F(\mathbf{L}) - Nf}\left(Nf + \frac{q}{p}\frac{\partial p}{\partial q}F(\mathbf{L})\right) + \sum_{\theta}\frac{\gamma_\theta L_\theta}{V'_\theta} = 0.$$

Therefore, the equilibrium mass of firms is implicitly given by

$$N^* = \frac{F(\mathbf{L})}{f}\frac{\sum_\theta \frac{\gamma_\theta L_\theta}{V'_\theta} - E\frac{q}{p}\frac{\partial p}{\partial q}}{\sum_\theta \frac{\gamma_\theta L_\theta}{V'_\theta} + E}. \tag{14}$$

Inspecting (14) reveals that increasing firms' market power on both the product market (the demand elasticity $-(q/p)(\partial p/\partial q)$ becomes smaller) and the labor market (γ_θ increases) results in *a higher number of firms under free entry*, the reason being that competition is relaxed on both markets.

Using (14) allows one to show how entry affects the structure of prices and wages. First, substituting (14) into (9), we obtain the free-entry equilibrium price:

$$p^* = \frac{1}{1 + \frac{q}{p}\frac{\partial p}{\partial q}} \frac{E + \sum_\theta \frac{\gamma_\theta L_\theta}{V'_\theta}}{F(\mathbf{L})}, \tag{15}$$

which now depends on the whole range of heterogeneity indices γ_θ. This price encapsulates a 'double markup' expressed through firms' relative markup on the product market and the weighted sum of heterogeneity indices on the labor markets. These two effects reinforce each other.

As for the free-entry equilibrium wage of a θ-worker, it is still given by (8) where p^* is now given by (15):

$$w^*_\theta = \left(E + \sum_t \frac{\gamma_t L_t}{V'_t} \right) \frac{F'_\theta}{F(\mathbf{L})} - \frac{\gamma_\theta}{V'_\theta}. \tag{16}$$

The expressions (14), (15) and (16) highlights how the two sources of market imperfection interact to determine the equilibrium number of firms, price and wages.

When firms' monopoly power on the product market increases, e.g., varieties become more differentiated, wages are depressed as long as the number of firms is given. However, the entry of new firms provides workers with a wider range of job opportunities, which tends to push wages upward. As a result, the impact of entry on wages is a priori undetermined.

In order to better understand what is going on, let us consider the benchmark case of perfectly competitive labor markets ($\gamma_\theta = 0$ for all θ). The corresponding equilibrium price and wages are given by

$$\hat{p} = \frac{\hat{E}}{F(\mathbf{L})} \frac{1}{1 + \frac{q}{p}\frac{\partial p}{\partial q}} \qquad \hat{w}_\theta = \frac{\hat{E}}{F(\mathbf{L})} F'_\theta, \tag{17}$$

where $\hat{E}$is the total expenditure on the differentiated good at the monopolistically competitive equilibrium.

Comparing (16) and (17) reveals that the free-entry equilibrium wage w^*_θ exceeds the competitive wage $\hat{w}_\theta$ if and only if

$$\left(\frac{E-\hat{E}}{\sum_t \gamma_t L_t / V_t'}+1\right)\frac{F_\theta'}{F\left(\mathbf{L}\right)} > \frac{\gamma_\theta / V_\theta'}{\sum_t \gamma_t L_t / V_t'}. \tag{18}$$

Since V_t' varies with the price level and the mass of varieties, which both depend on E,and thus on wages, we are unable to determine under which conditions (18) holds.

However, when consumers share the same CES preferences, things become much easier to interpret because all workers spend the same share of their earnings on the differentiated good, which makes the distribution of their earnings irrelevant for the determination of consumers' expenditures E. Specifically, we have

$$\frac{E-\hat{E}}{\mathbf{P}\sum_t \gamma_t L_t} = 0 \quad \text{ if } \mu_\theta = \mu,$$

so that (18) becomes

$$\frac{F_\theta'}{F\left(\mathbf{L}\right)} > \frac{\gamma_\theta}{\sum_t \gamma_t L_t}.$$

This inequality shows that the θ-workers earn a wage exceeding the competitive wage when the ratio of their marginal productivity to the average production of labor exceeds the ratio of their heterogeneity index to the average index. In particular, *workers with lower heterogeneity indices benefit from the presence of equally productive workers who have higher heterogeneity* as they extract more than their competitive wage. When workers differ in preferences (μ_θ), this effect is stronger when monopsonistic competition on the labor market redistributes workers' earnings towards those who have a higher μ_θ, the reason being that the total expenditure E increases and, eventually, exceeds $\hat{E}$. Thus, contrary to the conventional wisdom, we have:

Proposition 3. *Freeentry on the product market does not wash out the between-type redistributional effects of workers' heterogeneity.*

More generally, an increase in γ_θ gives rise to two opposite effects. First, it increases firms' monopsony power over the θ-workers, which allows the incumbent firms to pay them lower wages. Second, the incumbents make higher profits, which invites entry; this shifts upward the demand for labor, and thus push wages upward. As shown by differentiating (16) with respect to γ_θ, the former effect dominates the latter so that the θ-workers receive a lower pay. In contrast, an increase in γ_θ is always beneficial to all the other groups of workers because only the entry effect is at work. For example, observing a negative correlation between seniority and salary of university professors, Ransom (1993) argues along the same lines that "[i]ndividuals with high moving costs receive lower salary offers and have higher seniority than individuals with low moving costs." In the same vein, in a society where women would value hedonic job attributes more than men, such as time flexibility and home proximity, *women having the same productivity as men would earn lower wages*. What is more, the degree of gender discrimination varies with women's marital and family status, an empirical fact detailed by Polachek (2014), if we believe that married women with children value more hedonic job attributes than single women. Thus, our approach offers a theory of *differential discrimination*: discriminated workers belonging to the same group need not be equally discriminated.

Last, we show how our model can be used to shed new light on two important issues. First, workers' heterogeneity shows that employment density across space matters in an unsuspected way. Workers living in small cities operate in markets with few potential employers, so that workers must incur the costs of moving to another place if they want to earn higher wages. However, changing place typically involves various kinds of sunk costs, which makes these workers stickier. By contrast, workers living in large cities do not have to change place to face a large array of potential employers. This makes them more prone to change jobs. Consequently, workers having the same individual characteristics will earn higher wages in larger cities than in smaller cities because *firms have less monopsony power in thicker labor markets than in thinner ones* (Manning, 2010). Accordingly, even though it is well documented that the urban wage premium primarily stems from the presence of agglomeration economies at the city level (Rosenthal and Strange, 2004; Combes et al., 2012), we expect this effect to be exacerbated by a lower degree of monopsony power on a large urban labor market, in proportion that varies with technologies and workers' attitudes toward job and geographical mobility.

Second, according to the OECD (2012), the median labor share dropped from 66.1 percent in the early 1990s to 61.7 percent in the late 2000s.[1] Reinterpreting the input vector $\mathbf{l}(i)$ as the amounts of capital and labor needed to produce $q(i)+f$ units of the final good, our results provide a new perspective on the distribution of rents between employers and workers, which supplements those envisioned in the literature (Blanchard and Giavazzi, 2003). Indeed, it is natural to assume that hedonic attributes matter more to workers than to capital-owners. Under these circumstances, our results imply that *capital-owners capture a higher rent when workers are heterogeneous in taste for jobs, whereas sticky workers are hurt by the fact that capital-owners focus on the highest rate of return.* In particular, globalization would raise the rent accruing to capital as labor markets remain local or regional. Conversely, when capital is locked in specific locations such as the heavy or oil extraction industry, mobile workers exhibit a lower degree of heterogeneity than capital-owners, thus allowing the former to secure earnings that exceeds their competitive wages.

Since the above argument holds true for any two production factors, we have the next proposition.

Proposition 4. *Assume there is monopolistic competition on the product market and monopsonistic competition on the input markets. If there are two inputs, the input with the higher mobility across firms extracts more than its competitive earning, whereas the input with the lower mobility gains less.*

5 Concluding remarks

The analysis developed in this paper has several important implications. First of all, an elasticity of a firm's labor supply equal to 4 implies that on average workers accept a wage cut of 25 percent as a counterpart of the hedonic job attributes, while an elasticity of substitution across varieties equal to 7 implies that their marginal productivity is evaluated at 86 percent of the market price. In this event, the degree of exploitation of labor is far from being negligible. Second, a group of workers showing a high degree of attachment to specific job attributes is discriminated against compared to a group of workers who put a low weight on non-pecuniary characteristics. For example, in

[1] Jayadev (2007) finds a robust negative correlation between the degree of openness and the labor share in developed countries.

a society dominated by male chauvinist behaviors, women will earn less than men even when they both have the same productivity. Third, preference heterogeneity tends to exacerbate wage inequalities among workers' types. Last, institutions such as minimum wage rules or unions that push wages up more for lesser than higher skilled men (Card et al., 2004) reduce wage dispersion not only by raising the wage of the low-paid workers but also by indirectly decreasing those of the high-paid workers. A progressive income tax should play a similar role by making the high-paid workers less sensitive to gross wages.

Models of monopolistic competition has been extensively used in many economic fields. The tractability of our model, which combines monopsonistic competition and monopolistic competition, should permit its application to a wide range of issues. As a first extension, it seems natural to investigate how workers' search costs affect the labor market outcome. By analogy with what arises on the product market when consumers incur search costs, we may expect the following scenario to hold. First, the greater the workers' heterogeneity, the more workers search. This intensified search activity reduces firms' monopsony power and increases wages. However, once workers' heterogeneity is high enough, wages fall for the reasons discussed in this paper. This extension can then be grafted onto a setting in which a worker faces a positive probability of not to being hired by the firm she chooses. This should allow one to develop a new theory of unemployment based on imperfectly competitive markets.

Empirical evidence shows that firms differ in productivity. In this case, it is natural to expect the more productive firms to have higher sales, which requires a larger workforce. Being more productive, these firms can afford to pay higher wages to attract the additional workers they need. Since the high-productive workers value relatively more their wages than the other jobs' attributes, the more productive firms will enjoy a more productive labor force, thus magnifying their initial technological advantage through a positive assortative matching à la Sattinger (1993).

A word, in closing. We do not consider the approach developed in this paper as an alternative to other theories explaining the distribution of earnings. On the contrary, we see it as a "prediction-augmenting" theory that can be grafted onto others. Heterogeneity is pervasive in the real world and there is no apparent reason why labor markets should be immune.

Appendix. Proof of $dw^*/dN > 0$.

Plugging (10) into (11) and differentiating the resulting expression with respect to N yields

$$\frac{dw^*}{dN} = \mu \frac{f(\sigma-1)^2 HN + [H\mu(\sigma-1) - w^*(1-\mu)\sigma(L-Nf)](L-Nf\sigma)}{\sigma(\sigma-1)(1-\mu)N(L-Nf)^2}$$

It follows from (11) that

$$0 < w^* < \frac{\mu}{1-\mu}\frac{H}{L-Nf}\frac{\sigma-1}{\sigma}. \tag{A.1}$$

Assume, first, that $L - Nf\sigma > 0$. In this case, dw^*/dN decreases with w^*. Given (A.1), it is then readily verified that

$$\frac{dw^*}{dN} > \frac{\mu}{1-\mu}\frac{H}{(L-Nf)^2}\frac{\sigma-1}{\sigma}f > 0.$$

We now assume that $L - Nf\sigma < 0$. Therefore, dw^*/dN increases with w^*. Using (A.1), we obtain

$$\frac{dw^*}{dN} > \mu H \frac{\mu(L-Nf) + Nf(1-\mu)(\sigma-1)}{\sigma N(1-\mu)(L-Nf)^2} > 0.$$

References

[1] Anderson, J. and E. van Wincoop (2004) Trade costs. *Journal of Economic Literature* XLII, 691-751.

[2] Ashenfelter, O.C., H. Farber and M.R. Ransom (2010) Modern models of monopsony in labor markets: a brief survey. *Journal of Labor Economics* 28, 203-10.

[3] Ben-Akiva, M., N. Litinas and K. Tsunokawa (1985) Continuous spatial choice: The continuous logit model and distributions of trips and urban densities. *Transportation Research A* 19, 119-54.

[4] Bhaskar, V. and T. To (1999) Minimum wage laws for Ronald McDonald monopsonies: A theory of monopsonistic competition. *Economic Journal* 109, 190-203.

[5] Bhaskar, V. and T. To (2003) Oligopsony and the distribution of wages. *European Economic Review* 47, 371-99.

[6] Bhaskar, V., A. Manning and T. To (2002) Oligopsony and monopsonistic competition in labor markets. *Journal of Economic Perspectives* 16, 155-74.

[7] Black and Strahan (2001) The division of spoils: Rent-sharing and discrimination in a regulated industry. *American Economic Review* 91, 814-31.

[8] Blanchard, O.J. and N. Kiyotaki (1987) Monopolistic competition and the effects of aggregate demand. *American Economic Review* 77, 647-66.

[9] Blanchard, O. and F. Giavazzi (2003) Macroeconomic effects of regulation and deregulation in goods and labor markets. *Quarterly Journal of Economics* 118, 879-907.

[10] Boal, W.M. and M.R. Ransom (1997) Monopsony in the labor market. *Journal of Economic Literature* 35, 86-112.

[11] Card, D., T. Lemieux and W.C. Riddell (2004) Unions and wage inequality. *Journal of Labor Research* 25, 519-59.

[12] Dagsvik, J.K. (2002) Discrete choice in continuous time: Implications of an intertemporal version of the IIA property. *Econometrica* 70, 817-31.

[13] de la Rica, S., J.J. Dolado and R. Vegas (2010) Performance pay and the gender wage gap: Evidence from Spain. CEPR DP N°7936.

[14] Dixit, A.K. and J.E. Stiglitz (1977) Monopolistic competition and optimum product diversity. *American Economic Review* 67, 297-308.

[15] Docquier, F. and H. Rapoport (2012) Globalization, brain drain, and development. *Journal of Economic Literature* 50, 681-730.

[16] Eriksson, T. and N. Kristensen (2014) Wages or fringes? Some evidence on trade-offs and sorting. *Journal of Labor Economics* 32, 899-928.

[17] Hamilton, J., J.-F. Thisse and Y. Zenou (2000) Wage competition with heterogeneous workers and firms. *Journal of Labor Economics* 18, 453-72.

[18] Jayadev, A. (2007) Capital account openness and the labour share of income. *Cambridge Journal of Economics* 31, 423-43.

[19] Kim, S. (1989) Labor specialization and the extent of the market. *Journal of Political Economy* 97, 692-705.

[20] MacDonald, D.N. and M.O. Reynolds (1994) Are baseball players paid their marginal product? *Managerial and Decision Economics* 15, 443-57.

[21] Manning, A. (2003) *Monopsony in Motion. Imperfect Competition in Labor Markets*. Princeton, Princeton University Press.

[22] Manning, A. (2010) Agglomeration and monoposony power in labour markets. *Journal of Economic Geography* 10, 717-44.

[23] Marimon, R. and F. Zilibotti (1999) Unemployment versus mismatch of talents: Reconsidering unemployment benefits. *Economic Journal* 109, 266-91.

[24] Moretti, E. (2012) *The New Geography of Jobs*. Boston, Houghton Mifflin Harcourt.

[25] OECD (2012) *OECD Employment Outlook*. Paris, OECD.

[26] Pigou, A.C. (1924) *The Economics of Welfare*. 2nd ed., London, Macmillan.

[27] Polachek, S.W. (2014) Equal pay legislation and the gender wage gap. *IZA World of Labor 16*.

[28] Ransom, M.R. (1993) Seniority and monopsony in the academic labor market. *American Economic Review* 83, 221-33.

[29] Robinson, J. (1933) *The Theory of Imperfect Competition*. London, Macmillan.

[30] Rosenthal, S.S. and W.C. Strange (2004) Evidence on the nature and sources of agglomeration economies. In: J.V. Henderson and J.-F. Thisse (eds.), *Handbook of Regional and Urban Economics. Cities and Geography*, pp. 2119-71. Amsterdam, North-Holland.

[31] Salop, S.C. (1979) Monopolistic competition with an outside good. *Bell Journal of Economics* 10, 141-56.

[32] Sattinger, M. (1993) Assignment models of the distribution of earnings. *Journal of Economic Literature* 31, 831-80.