

IZA DP No. 5136

The Distribution of Earnings under Monopsonistic/polistic Competition

Jacques-François Thisse
Eric Toulemonde

August 2010

The Distribution of Earnings under Monopsonistic/polistic Competition

Jacques-François Thisse

*CORE, Université catholique de Louvain,
Université du Luxembourg and CEPR*

Eric Toulemonde

*FUNDP-University of Namur,
CORE and IZA*

Discussion Paper No. 5136
August 2010

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

The Distribution of Earnings under Monopsonistic/polistic Competition*

Recent empirical contributions in labor economics suggest that individual firms face upward sloping labor supplies. We rationalize this by assuming that idiosyncratic non-pecuniary conditions interact with money wages in workers' decisions to work for specific firms. Likewise, firms supply differentiated goods in response to differences in consumer tastes. Hence, firms are price-makers and wage-setters. By combining monopolistic and monopsonistic competition, our setting encapsulates general equilibrium interactions between the two markets. The equilibrium involves double exploitation of labor. Compared to the competitive outcome, the high-productive workers are overpaid under free entry, whereas the low-productive workers are underpaid. In the same vein, capital-owners receive a premium, whereas workers are exploited.

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

Keywords: worker heterogeneity, monopsonistic competition, monopolistic competition, labor exploitation, wage dispersion

Corresponding author:

Eric Toulemonde
FUNDP
Faculté des Sciences Economiques, Sociales et de Gestion
8, Rempart de la Vierge
B-5000 Namur
Belgium
E-mail: eric.toulemonde@fundp.ac.be

* This research was undertaken under the support of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by the authors.

1 Introduction

The economics of labor markets has been, and to a large extent is still, dominated by the competitive paradigm. This means that, by slightly decreasing the wage it pays a firm would lose all its workers. Surveying recent empirical contributions that aims to measure the extent of monopsony power in labor markets, Ashenfelter et al. (2010) conclude that firms face labor supplies that are *not* infinitely elastic. This fact may be rationalized by assuming that idiosyncratic non-pecuniary conditions strongly interact with the role of money wages in workers' decisions to accept offers made by specific firms. Empirical work confirms the idea that the standard assumption of identical firms glosses over many important aspects of reality. In particular, firms' behavior build on their core competencies and corporate culture because these ones can hardly be imitated, thus making them inherently heterogeneous (Berger, 2005). As a consequence, heterogeneous workers view jobs offered by different firms as bundles of non-wage job attributes which provide them with more or less satisfaction, very much as heterogeneous consumers perceive differentiated products as different bundles of non-price product characteristics (Rosen, 2002). This agrees with Hamermesh (1977) who observes that the gap between actual wages and competitive wages is positively correlated with various measures of job satisfaction. Likewise, it is well documented in industrial organization that producers supply differentiated goods in response to differences in consumer tastes. In such a context, firms are endowed with market power that allow them to be both price-makers and wage-setters (Manning, 2003). As will be shown in this paper, blending imperfections on the product and labor markets within a unified framework yields new and insightful results about the distribution of earnings, which all agree with empirical evidence.

More precisely, we distinguish two market environments. In the first place, when the number of firms is exogenous, we show that the equilibrium involves *double exploitation of workers*, meaning that workers are paid below their marginal value product for two distinct reasons. The former stems from the heterogeneity of workers that endows firms with monopsony power. Because each worker has a most-preferred employer, firms may set a lower wage while attracting their captive labor pool. The latter results from the fact that firms are price-makers on the product market because they sell a differentiated good. In this case, they evaluate workers' marginal productivity at the marginal revenue, which is lower than the market price. Hence, workers' marginal productivity is undervalued. All in all, once it is recognized that both the product and labor markets operate under imperfect competition, the market yields an inefficient outcome that involves an income transfer from the workers to the firms. It is worth stressing that labor exploitation is here the involuntary consequences of a myriad of individual behaviors made by firms in an environment in which they behave non-cooperatively and where any single firm is negligible to the market.

Furthermore, we show that the equilibrium is such that *the wage gap exceeds the productivity gap* between any two different types of workers. Indeed, a worker's genuine wage is equal to her actual wage plus the monetary evaluation of the hedonic attributes of her job. Since the relative value of the hedonic attributes is lower for the high-wage workers than for the low-wage workers, the high-wage workers are more sensitive to wage differences than the low-wage workers. Hence, the latter are more exploited than the former, which means that workers' heterogeneity magnifies productivity differences. Last, even when workers have the same productivity, wage dispersion may arise because workers need not have the same

preferences. Firms then exploit workers' idiosyncrasies to set different wages. For instance, de la Rica et al. (2010) find that monopsonistic features, which could be related to women's lower labor mobility due to housework, explain the gender wage gap in Spain.

In the second place, when the number of firms is endogenous and determined by free entry and exit, the income transfer from workers to firms vanishes because profits are zero. This does not mean, however, that workers' exploitation is washed out by free entry. Labor exploitation still matters through income transfers across different types of workers. More precisely, the double exploitation gives rise to general equilibrium effects stemming from the entry of new firms, which affects the intensity of competition in the product and labor markets. The above-mentioned magnification of productivity differences also holds at free entry. Since profits are zero, this effect expresses itself through an income transfer from the low-productive workers to the high-productive workers. More precisely, we show that, compared to the competitive outcome, *the high-productive workers are overpaid, whereas the low-productive workers are underpaid*. This magnification effect also implies that a technological shock beneficial to the high-skilled workers exacerbates wage inequality between high- and low-skilled workers. The heterogeneity of workers' attitudes toward job offers is thus likely to be part of the explanation for the growing wage inequality between high- and low-skilled workers.

In the same vein, workers who value less jobs' hedonic attributes gain at the expense of those for whom these attributes are more important, even when they have the same productivity. Therefore, if a group of workers value more such attributes than another, they will be discriminated against even if they share identical observable characteristics in every other aspect. Likewise, if the production factors are capital and labor, we find it natural to suppose that hedonic attributes matter more to workers than to capital-owners. Hence, *capital-owners would receive a premium, whereas workers would be underpaid*.

To prove those results, we need a general equilibrium model that accounts for imperfect competition on both goods' and labor markets. Although a comprehensive general equilibrium model with strategic interactions has so far been out of reach and is likely to remain so for a long time, it is possible to gain insights from specific models. To be precise, we model the product market using the CES model of monopolistic competition of the Dixit-Stiglitz genre, which permits one to work with imperfect competition and increasing returns on the product market. As for the labor market, because heterogeneous 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 as the realization of a random variable and, therefore, their heterogeneous response to a wage cut (Anderson et al., 1992). Specifically, we assume that workers' heterogeneity is captured by the logit model.

We acknowledge the fact that our model is very specific. That said, we want to emphasize that our modeling strategy has several merits. First, the CES is the workhorse of many economic fields addressing imperfect competition, while the logit is its natural counterpart. Indeed, they both obey the same rules, apart from the fact that the former deals with divisible consumption choices and the latter with indivisible job choices (Anderson et al., 1992). Second, our main results are sufficiently intuitive to conjecture that they hold true in more general frameworks. Third, our model is amenable to the data because both the CES and the logit can be tested by means of very efficient econometric techniques. The flexibility of these two models should also permit to work with more general frameworks in empirical

studies.

Last, as observed by Bhaskar et al. (2002) and Manning (2003), monopsony power on the labor market does not imply the existence of a single employer. It means only that a firm does not lose all its workers when slightly decreasing its wage. This observation has led several authors to build on Salop (1979) to model heterogeneous firms competing to attract heterogeneous workers. In such a setting, the nature of competition is oligopsonistic because the competitive forces come from the two adjacent firms (Kim, 1989; Bhaskar and To, 1999, 2003; Marimon and Zilibotti, 1999; Hamilton et al., 2000). The introduction of strategic considerations in the labor market makes the analysis much more involved, thus leading these authors to consider the product side as perfectly competitive. They fail to capture, therefore, any type of interaction that could stem from imperfections on the product and labor markets.

In contrast, by using *monopolistic* and *monopsonistic* competition, our setting encapsulates general equilibrium interactions *between* the two markets. It also captures the main aspect of imperfect competition, while avoiding the technical difficulties implied by strategic interactions *within* each market. Moreover, our modeling approach leads to a tractable framework, which translates into simple and neat results. To the best of our knowledge, our paper is the first one that addresses imperfect competition in the labor market by means of monopsonistic competition in the tradition of Dixit and Stiglitz (1977).¹

The paper proceeds as follows. In Section 2, we present the various assumptions made to describe consumption and production. In Section 3, we characterize the equilibrium when the number of firms is exogenous, while the subsequent section considers the case of free entry and exit. The last section summarizes our main policy implications and discusses possible applications.

2 The model

2.1 Technology

We consider an economy endowed with one sector and Θ types of labor. There are L_θ workers of type θ and each worker is endowed with one unit of her type of labor. The total population is denoted by $L = \sum_\theta L_\theta$, while $\mathbf{L} \equiv (L_1, \dots, L_\Theta)$. 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 under increasing returns and monopolistic competition; it is made available as a continuum of varieties of mass N . Firm i supplies variety $i \in [0, N]$ and hires $\ell_\theta(i)$ workers of type $\theta = 1, \dots, \Theta$. This firm's production function is given by a linear homogeneous function $F[\mathbf{l}(i)]$, where $\mathbf{l}(i) \equiv (\ell_1(i), \dots, \ell_\Theta(i))$. Its output $F[\mathbf{l}(i)]$ is split between the quantity $q(i)$ offered to consumers and the fixed requirement f needed to undertake production. In other words, we have:

$$q(i) = F[\mathbf{l}(i)] - f. \quad (1)$$

¹To (2009) uses a CES function to describe the aggregate labor market, but he does not address the individual job choices made by workers.

When there is a single type of workers, the unit of labor is chosen for the production function to be written as follows:

$$q(i) = \ell(i) - f.$$

Note that our model need not be confined to different types of labor as inputs. It displays enough versatility to deal with different production factors such as capital and labor.

2.2 Workers

Preferences. Workers have a love for variety and share the same quasi-linear preferences nesting a CES-subutility:

$$U(M, h) = \alpha \ln M + h \quad \alpha > 0 \quad (2)$$

where h is the individual consumption of the homogeneous good and M the composite good given by

$$M \equiv \left[\int_0^N (x(i))^{\frac{\sigma-1}{\sigma}} di \right]^{\frac{\sigma}{\sigma-1}}$$

in which $x(i)$ is the individual consumption of variety i of the differentiated good, while $\sigma > 1$ is the elasticity of substitution across varieties. Econometric estimations suggest that σ varies from 5 to 10 (Anderson and van Wincoop, 2004). Last, in (2), $\alpha > 0$ is a measure of the desirability of the differentiated good relatively to the homogeneous good.

A θ -worker hired by firm i earns a nominal wage $w_\theta(i)$ and has a budget constraint given by

$$PM + h = w_\theta(i) + \frac{H}{L} + \frac{1}{L} \int_0^N \pi(i) di$$

where

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

is the price index of the differentiated good and $\pi(i)$ the profit made by firm i (see below for more details). Assuming that the initial endowment H is sufficiently large for the consumption of the homogeneous good to be positive in equilibrium, it is readily verified that a worker's expenditures on the differentiated good is such that $PM = \alpha$. As expected, for a given P , the consumption of this good increases with α . Furthermore, the individual inverse demand for variety i is:

$$p(i) = (\alpha/x(i))^{\frac{1}{\sigma}} P^{\frac{\sigma-1}{\sigma}}. \quad (3)$$

Workers' heterogeneity. Workers are free to choose the firm they want to work for. As discussed in the introduction, they are heterogenous in their perception of the non-wage attributes associated with a particular firm. Formally, the indirect utility of a θ -worker employed in firm i is given by

$$V_\theta(i) = \alpha (\ln \alpha - 1) - \alpha \ln P + \frac{H}{L} + \frac{1}{L} \int_0^N \pi(i) di + w_\theta(i) + \varepsilon_\theta(i)$$

where the quality of her match with firm i is given by the realization of the 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 indirect utility, that is, the *hedonic wage* given by

$$\max_i [w_\theta(i) + \varepsilon_\theta(i)]$$

which depends on the wages set by firms and the levels of the worker's match. We assume that the random variables $\varepsilon_\theta(i)$ are independently and identically distributed according to the Gumbel distribution with zero mean.² This implies that the probability she chooses to work in firm i is given by the *continuous logit* (McFadden, 1976; Ben-Akiva et al., 1985; Dagsvik, 2002):

$$\mathbb{P}_\theta(i) = \frac{\exp \frac{w_\theta(i)}{\nu_\theta}}{\int_0^N \exp \frac{w_\theta(j)}{\nu_\theta} dj} \quad (4)$$

where ν_θ stands for the standard-deviation of $\varepsilon_\theta(i)$ (up to the numerical factor $\pi/\sqrt{6}$). Such a modeling strategy allows one to account for a population of θ -workers exhibiting heterogeneous tastes about firms/jobs since the probability that $\varepsilon_\theta(i) = \varepsilon_\theta(j)$ is zero. In (4), ν_θ is an index that captures the diversity of preferences across θ -workers who react differently to the same wage schedule. Alternatively, ν_θ may be interpreted as an inverse measure of θ -workers' *inter-firm mobility*: a larger ν_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 ν_θ is small enough for the expressions derived below to be positive.

As shown by Ben-Akiva et al. (1985), the expected hedonic wage is equal to

$$\nu_\theta \ln \int_0^N \exp \left[\frac{w_\theta(i)}{\nu_\theta} \right] di.$$

When nominal wages are equalized across firms, this expression becomes $w_\theta + \nu_\theta \ln N$, which increases at a decreasing rate with the mass of firms. Workers' heterogeneity thus translates into a preference for job variety.

2.3 Firms

Since firms sell differentiated varieties, each firm can freely choose the price of its variety. Since workers have heterogeneous preferences across employers, each firm is free to set its wage for each type of labor. Specifically, when firm i sets the wage $w_\theta(i)$, it attracts

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

workers of type θ . As a consequence, firms face *monopolistic competition* on the product market and *monopsonistic competition* on the labor market.

²Because the support of the Gumbel distribution that generates the logit is the real line, a worker's highest hedonic wage could be negative. However, we may disregard this issue because each worker faces a continuum of firms.

Firm i chooses the price $p(i)$ and the wages $w_\theta(i)$ to maximize its profits given by

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

subject to (1) and (3).

Market clearing for variety i means that $q(i) = Lx(i)$. Substituting (3) for $p(i)$ and (1) for $x(i)$ yields the following expression for $\pi(i)$:

$$\pi(i) = (\alpha L)^{\frac{1}{\sigma}} \{ [F(\mathbf{1}(i)) - f] P \}^{\frac{\sigma-1}{\sigma}} - \sum_{\theta} w_{\theta}(i) \ell_{\theta}(i) \quad (5)$$

where $\ell_{\theta}(i) = L_{\theta} \mathbb{P}_{\theta}(i)$. Hence, solving firm i 's profit-maximization program amounts to maximizing (5) with respect to $w_{\theta}(i)$.

In order to disentangle the various effects at work, it is both relevant and convenient to distinguish between what we call a *short-run equilibrium*, in which the mass N of firms is fixed, and a *long-run equilibrium* in which the mass of firms is endogenously determined through free entry and exit.

3 The short-run equilibrium

3.1 Wage and price

Since firm i is negligible to the labor market, it accurately treats the denominator of (4) parametrically. Hence,

$$\frac{d\ell_{\theta}(i)}{dw_{\theta}(i)} = \frac{1}{\nu_{\theta}} \ell_{\theta}(i). \quad (6)$$

so that the elasticity of firm i 's labor supply is

$$e_{\theta}(i) = \frac{w_{\theta}(i)}{\nu_{\theta}}.$$

Hence, as long as θ -workers are heterogeneous ($\nu_{\theta} > 0$), firm i faces a supply curve with a finite elasticity: *ceteris paribus*, the more heterogeneous the θ -workers, the smaller this elasticity. In contrast, the supply curve is infinitely elastic provided that θ -workers are homogeneous, that is, they care only about their wage ($\nu_{\theta} = 0$). Furthermore, this firm's labor supply becomes more (resp., less) elastic when the θ -workers earn a higher (resp., lower) wage, thus suggesting that high-productive workers are more responsive to wage differences than low-productive workers.

Differentiating (5) with respect to $w_{\theta}(i)$ and using (6), we obtain the following equilibrium wage:

$$w_{\theta}^{*}(i) = \frac{\sigma - 1}{\sigma} P \left\{ \frac{\alpha L}{[F(\mathbf{1}(i)) - f] P} \right\}^{\frac{1}{\sigma}} F'_{\theta} - \nu_{\theta} \quad (7)$$

where F'_{θ} denotes the derivative of F with respect to $\ell_{\theta}(i)$. In this expression, the first term stands for the marginal revenue generated by a θ -worker. The second term captures

the fact that workers' heterogeneity endows firms with monopsony power. Since $\nu_\theta > 0$, firms are able to pay the θ -workers a wage smaller than the marginal revenue they generate. In other words, the degree of workers' heterogeneity determines the degree of exploitation. Note, however, that θ -workers are paid the same wage although they face different matching values.

Given the symmetry of our setting, we find it natural to focus on the symmetric outcome:

$$\ell_\theta(i) = L_\theta/N \quad x(i) = [F(\mathbf{L}) - Nf]/NL \quad p(i) = p.$$

Substituting $x(i)$ and $P \equiv N^{\frac{-1}{\sigma-1}}p$ into (3) yields the equilibrium price:

$$p^*(N) = \frac{\alpha L}{F(\mathbf{L}) - Nf} \quad (8)$$

This expression has several implications. First, when the labor force L remains constant, the entry of new firms gives rise to two opposite effects. A larger number of firms reduces the output of each firm ($F(\mathbf{L}/N) - f$), which fosters a higher price. In addition, a larger number of firms also makes competition tougher, thus pushing the market price downward. However, it is well known that the latter effect vanishes under the CES. Consequently, the former effect is the only one at work. Indeed, inspecting (8) reveals that p^* increases with the mass N of firms.

Second, when the labor force increases homothetically, the equilibrium price goes down. Indeed, although the total expenditure (αL) and the total production ($F(\mathbf{L})$) increase at the same rate, the total production available for consumption ($F(\mathbf{L}) - Nf$) increases at an even higher rate, which yields a lower price. Last, observe that (8) is independent from the elasticity of substitution across varieties. This is because the quasi-linearity of preferences implies that the amount spent on the differentiated product is equal to αL , while the denominator $F(\mathbf{L}) - Nf$ is constant when N is fixed.

The equilibrium wage is given by

$$w_\theta^*(N) = \frac{\sigma - 1}{\sigma} p^*(N) F'_\theta - v_\theta \quad (9)$$

which after replacement yields

$$w_\theta^*(N) = \frac{\sigma - 1}{\sigma} \frac{\alpha L}{F(\mathbf{L}) - Nf} F'_\theta - v_\theta. \quad (10)$$

3.2 Labor exploitation

As shown by (9), the equilibrium wage of θ -workers differs from their marginal value product, pF'_θ . The exploitation of workers has two sources here. First, firms use their monopoly power on the product market to set a lower wage equal to the marginal value product times the inverse of the relative markup: $(\sigma - 1)/\sigma < 1$. If the M -good were homogeneous, this markup would be equal to 1. An elasticity of substitution varying from 5 to 10 suggests a first exploitation rate of about 10 to 20 percent of the marginal value product. Second, since θ -workers exhibit diversity in their preferences for employers, firms use their monopsony

power on the labor market to reduce the wage by v_θ , which is itself given by the share $1/e_\theta$ of the wage. As noticed by Boal and Ransom (1997), Pigou (1924) used the ratio $1/e_\theta = \nu_\theta/w_\theta$ 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. Consequently, we may conclude that *there is "double exploitation" of labor*, that is, an income transfer away from workers stemming from the fact that both the product and labor markets are imperfectly competitive.

In order to shed light on this double exploitation, assume that there is a single type of labor ($\Theta = 1$) with $F(\mathbf{L}) = L$ and constant returns ($f = 0$). We first determine the equilibrium wage in the special case of homogeneous workers ($v = 0$). Using (10), we obtain

$$w^* = \frac{\sigma - 1}{\sigma} \alpha. \quad (11)$$

Since individual expenditures on the differentiated product is α , the total revenue in the M -sector is equal to αL . To begin with, assume that firms price at marginal cost. Under constant returns, equilibrium profits are zero. Therefore, the *socially optimal wage* is given by $w = \alpha$. Assume now that firms are free to choose their price on a monopolistically competitive market. In order to hire workers, they enter into a bidding process in which they increase their offer up to the point where their bid equals their marginal revenue. Since this one is smaller than the market price, the resulting equilibrium wage must be smaller than α . This is precisely what (11) states: when monopolistic competition prevails on the product market, the equilibrium wage is equal to the socially optimal wage times the inverse relative markup. In other words, the equilibrium wage decreases with the degree of firms' monopoly power on the product market. Even in the absence of imperfections on the labor market, imperfect competition on the product market translates into a wage smaller than the socially optimal wage because firms strives to produce less and, consequently, to hire less workers.

Consider now the reverse thought experiment in which firms selling a homogeneous good operate under perfect competition, while workers are heterogeneous. In this case, since $p^* = \alpha$, (10) becomes

$$w^* = \alpha - v.$$

Firms now build their monopsony power on the diversity of workers' preferences. This allows them to extract from their workers a premium v , which yields again an equilibrium wage smaller than the socially optimal wage.

It remains to understand the role of increasing returns in the determination of the equilibrium wage. Considering again the case of a single type of heterogeneous workers, we obtain

$$w^* = \alpha \frac{\sigma - 1}{\sigma} \frac{L}{L - Nf} - v.$$

This expression shows that the presence of increasing returns implies that the relative desirability of the differentiated product, measured by $\alpha L/(L - Nf)$, rises with f . This now leads firms to pay a higher wage to their workers because the total mass of workers net of fixed requirements, $L - Nf$, is lower than under constant returns.

That said, it should be clear that (10) reflects the presence of those various effects encapsulated within a setting that involves several types of heterogeneous workers. To illustrate, consider now two different types of labor (k and l) having the same diversity index ($\nu_k = \nu_l = \nu$). Then, we have

$$w_k = \frac{F'_k}{F'_l} w_l + \nu \left(\frac{F'_k}{F'_l} - 1 \right).$$

As expected, the workers with the higher marginal product earn more than those with the lower marginal product. Indeed, $w_k > w_l$ holds if and only if $F'_k/F'_l > 1$. Furthermore, the wage gap exceeds the productivity gap since the ratio w_k/w_l increases with ν as long as $F'_k/F'_l > 1$. In other words, *heterogeneity makes the low-productivity workers relatively worse off with respect to the high-productivity workers*. This is because the relative value of the match 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. MacDonald and Reynolds (1994) 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. They also 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. A priori, we do not rule out the possibility that a large number of high-skilled workers may have a lower marginal product than a small number of low-skilled.

When the two types of labor have different diversity indices ($\nu_k \neq \nu_l$), the foregoing expression becomes

$$w_k = \frac{F'_k}{F'_l} w_l + \nu_k \left(\frac{\nu_l}{\nu_k} \frac{F'_k}{F'_l} - 1 \right).$$

What matters is now workers' marginal product *weighted* by their diversity. For instance, if the high-productivity workers value more the quality of the match than the low-productivity ones ($\nu_k > \nu_l$), they might be willing to take a job with a lower pay because they enjoy a higher non-monetary compensation.

4 The long-run equilibrium

4.1 Wage dispersion

Substituting (8) into P , the zero-profit condition becomes

$$\pi(i) = \frac{1}{N} \left(\alpha L - \sum_{\theta} w_{\theta} L_{\theta} \right) = 0.$$

Plugging (10) into this expression and solving with respect to N , we obtain

$$N^* = \left(1 - \frac{\alpha}{\alpha + \bar{\nu}} \frac{\sigma - 1}{\sigma} \right) \frac{F(\mathbf{L})}{f} > 0 \quad (12)$$

where

$$\bar{\nu} = \sum \frac{L_\theta}{L} \nu_\theta$$

is the average match. Inspecting (12) reveals that more market power on the product (σ decreases) and labor ($\bar{\nu}$ increases) markets leads to a higher number of firms under free entry.

Using N^* allows one to show how entry affects the structure of prices and wages. First, substituting (12) into (8), we obtain the long-run equilibrium price

$$p^* = \frac{\sigma}{\sigma - 1} \frac{\alpha + \bar{\nu}}{F(\mathbf{L})/L}.$$

This price encapsulates a “double premium” expressed through the relative markup and the average match, which reinforce each other. Specifically, unlike what we observe under a fixed number of firms, the sources of imperfection in both markets impact on the long-run equilibrium price through the entry or exit of firms.

Second, the long-run equilibrium wage of a θ -worker is given by

$$w_\theta^* = \frac{\alpha + \bar{\nu}}{F(\mathbf{L})/L} F'_\theta - \nu_\theta. \quad (13)$$

Unlike the short-run wage, the long-run wage is not affected by firms’ monopoly power on the product market. In contrast, monopsony power matters in a way that needs clarification.

(i) When workers’ groups are homogeneous ($\nu_\theta = 0$ for all θ), the competitive equilibrium wages are

$$w_\theta^c = \alpha \frac{F'_\theta}{F(\mathbf{L})/L}. \quad (14)$$

(ii) When there is a single type of heterogeneous workers, we have $F(L) = LF'(L)$ and $\bar{\nu} = \nu_\theta$. In this case, (13) boils down to the socially optimal wage:

$$w^* = \alpha = w^c.$$

Under free entry, firms use their monopoly power in the product market to generate markups that are just sufficient to cover their fixed costs. In contrast, firms are not able to exploit workers’ heterogeneity because entry washes out their monopsony power in the labor market. This shows how peculiar may be a modeling strategy relying on the single type of labor.

(iii) When there are several groups of heterogeneous workers who are equally productive in the sense that $LF'_\theta = F(\mathbf{L})$, we have

$$w_\theta^* = \alpha + \bar{\nu} - \nu_\theta.$$

Hence, *there is wage dispersion even when workers have the same productivity*. More precisely, the more (resp., less) heterogeneous workers earn a wage smaller (resp., higher) than the average wage α . This is because workers who value specific firms are willing to work for a lower pay, whereas employers have to pay higher wages to attract workers who are more or less indifferent across firms.

To sum up, under free entry and exit, *both workers’ heterogeneity and several types of labor are needed for wage dispersion to arise around the competitive wage*.

4.2 The distribution of earnings

Comparing (13) and (14), we obtain

$$w_\theta^* > w_\theta^c \iff \frac{F'_\theta}{F(\mathbf{L})/L} > \frac{\nu_\theta}{\bar{\nu}}.$$

Therefore, the θ -workers earn a wage higher than the competitive wage once the ratio of their marginal product to the average production of labor exceeds the ratio of their diversity index to the average index. In particular, one may expect productive workers with a low diversity index to benefit from the presence of workers with high diversity indices to extract more than the competitive wage. This has a worth-stressing implication in the special, but relevant, case of two production factors, labor and capital. Indeed, capital-owners are better-off when workers are heterogeneous, whereas workers are hurt by the fact that typical capital-owners seek the highest rate of return. This holds true as long as capital-owners display a diversity index smaller than workers'. Conversely, when capital is locked in specific locations such as the heavy industry, mobile workers exhibit a lower degree of diversity that allow them to secure earnings exceeding their competitive wages. In a nutshell, *by extracting more than its competitive earning, the production factor with the higher mobility across firms gains at the expense of the production factor with the lower mobility.*

More generally, an increase in ν_θ gives rise to two contrasting effects. First, it increases firms' monopsony power over the θ -workers, which allow firms to pay them a lower wage. Second, firms make higher profits, thus triggering the entry of new firms. This shifts upward the demand for labor, thus raising workers' earnings. As shown by differentiating (13) with respect to ν_θ , the increase in ν_θ is always detrimental to the θ -workers who get more exploited by their employers. In contrast, this increase is always beneficial to all the other groups of workers because it promotes entry only. For example, in a society where women would value more than men hedonic attributes, such as time flexibility and home proximity, *women having the same productivity as men would earn lower wages.* In addition, observing a negative correlation between seniority and salary of university professors, Ransom (1993) argues that “[i]ndividuals with high moving costs receive lower salary offers and have higher seniority than individuals with low moving costs.” This explanation concurs with our results.

If the θ -workers have a diversity index identical to the average index ($\nu_\theta = \bar{\nu}$), these workers benefit from monopsonistic competition if they belong to a high productivity group ($F'_\theta > F(\mathbf{L})/L$) and suffer from it otherwise. To illustrate, assume that all types of workers are perfect substitutes. In this case, the production function is given by $F(\mathbf{L}) = \Sigma a_\theta L_\theta$ so that

$$(\alpha + \bar{\nu}) \frac{a_\theta}{\Sigma a_\theta L_\theta / L} - \nu_\theta = w_\theta^* > w_\theta^c = \alpha \frac{a_\theta}{\Sigma a_\theta L_\theta / L} \iff \frac{a_\theta}{\Sigma a_\theta L_\theta / L} > \frac{\nu_\theta}{\bar{\nu}}. \quad (15)$$

In words, if the productivity of a group of workers differs from the average productivity more (resp., less) than their diversity index differs from the average index, then they earn more (resp., less) than the wage they would get on a competitive labor market. In the limit, if all workers have the same diversity index, the wage of the more (resp., less) productive workers is higher (resp., lower) than their respective competitive wage. This is so because the relative value of the hedonic attributes is higher for low productive workers than for high productive workers. The presence of workers with a *relatively* low inter-firm mobility allows firms to

pay them lower wages, thus inviting entry. As a consequence, there is more competition on the labor markets, especially for the workers who value *relatively* less the hedonic attributes of jobs, i.e. those with a high productivity. Accordingly, *the high-productive workers are overpaid, whereas the low-productive workers are underpaid, thus implying an implicit income transfer from the latter to the former since firms' profits are zero.* This concurs with MacDonald and Reynolds (1994) who found senior pitchers to be overpaid relative to their marginal product.

Furthermore, (15) reveals that, everything else being equal, the discrepancy between the equilibrium and the competitive wage ($w_\theta^* - w_\theta^c$) rises as the level of θ -workers' productivity increases. Therefore, the heterogeneity of workers' preferences magnifies the advantage of being more productive and the disadvantage of being less productive. This pecuniary externality triggers a rat race in that it provides all workers with incentives to acquire a level of human capital exceeding the level that would be socially optimal with homogeneous workers. The above results have another noticeable implication. Since academics are likely to display fairly heterogeneous preferences, universities will pay disproportionately high salaries to the super-stars, while underpaying the others.

More generally, if, for whatever reason, the marginal product of a group k evaluated at $F(\mathbf{L})$ rises while the marginal product of the other groups remains the same, the total production increases. Inspecting (13) shows that the wage of the k -workers rises, whereas the wages of the other groups fall. In particular, consider two types of labor, the skilled and the unskilled. If the skilled benefit from a positive technological shock that make them more productive, they earn a higher wage whereas the unskilled experience an absolute wage decrease. Hence, our setting provides a rationale for *the growing wage inequality between high- and low-skilled workers observed during the last 30 years in many developed countries.*

5 Concluding remarks

The fact that both sides of the product and labor markets are heterogeneous has several important policy 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% 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% of the market price. Second, a group of workers showing a high degree of attachment to specific job attributes are discriminated against compared to a group of workers who put a low weight on non-wage characteristics. For example, in a society dominated by male chauvinist behaviors, women will earn less than men even when they both have the same productivity. Third, if capital exhibits more inter-firm mobility than labor, capital-owners capture a rent at the expense of workers. How big is this rent is an empirical question that cannot be addressed here. Fourth, since high-productive workers are overpaid, individuals invest too much in human capital, thus leading to excessively high expenditures in education. Fifth, 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. Tax progressivity should play a

similar role by reducing more the elasticity of firms' supply of high-paid workers than that of the low-paid workers.

The CES model of monopolistic competition has been extensively used in many economic fields (Matsuyama, 1995). The tractability of our model, which nests monopsonistic competition into a standard model of monopolistic competition, should permit its application to a wide range of issues. In particular, given the extensive use of the CES in international trade, it seems natural to investigate how trade barriers and monopsony power interact to determine the level of trade flows. Moreover, an exporting firm must hire a larger number of workers to supply its bigger outlet. To do this, it must set higher wages than the non-exporting firms. In other words, exporting generates additional costs. To obviate the increase in the cost of labor, enterprises facing a growing demand may choose to set up production facilities in several regions or countries, even in the absence of trade costs. Another line of research that seems worth pursuing is to study how the size and structure of the labor force affect the price and wage levels and to compare the corresponding results with those obtained in more standard approaches.

Last, we have assumed that firms are homogeneous. However, evidence is mounting that they are heterogeneous. In this case, it is natural to expect the more productive firms to have higher sales, which requires a larger labor force. 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 monetary 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 composition effect.

References

- [1] Anderson, S.P., A. de Palma and J.-F. Thisse (1992) *Discrete Choice Theory of Product Differentiation*. Cambridge, MA, The MIT Press.
- [2] Anderson, J. and E. van Wincoop (2004) Trade costs. *Journal of Economic Literature* XLII, 691-751.
- [3] 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, 203-210.
- [4] 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, 119-154.
- [5] Berger, S. (2005) *How We Compete: What Companies Around the World Are Doing to Make It in the Global Economy*. New York, Doubleday.
- [6] Bhaskar, V. and T. To (1999) Minimum wage laws for Ronald McDonald monopsonies: a theory of monopsonistic competition. *Economic Journal* 109, 190-203.
- [7] Bhaskar, V. and T. To (2003) Oligopsony and the distribution of wages. *European Economic Review* 47, 371-399.

- [8] Bhaskar, V., A. Manning and T. To (2002) Oligopsony and monopsonistic competition in labor markets. *Journal of Economic Perspectives* 16, 155-174.
- [9] Boal, W.M. and M.R. Ransom (1997) Monopsony in the labor market. *Journal of Economic Literature* 35, 86-112.
- [10] Card, D., T. Lemieux and W.C. Riddell (2004) Unions and wage inequality. *Journal of Labor Research* 25, 519-559.
- [11] Dagsvik, J.K. (2002) Discrete choice in continuous time: implications of an intertemporal version of the IIA property. *Econometrica* 70, 817-831.
- [12] 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.
- [13] Dixit, A.K. and J.E. Stiglitz (1977) Monopolistic competition and optimum product diversity. *American Economic Review* 67, 297-308.
- [14] Hamermesh, D. (1977) Economic aspects of job satisfaction. In O. Ashenfelter and W. Oates (eds.), *Essays in Labor Market Analysis*. New York, Wiley.
- [15] Hamilton, J., J.-F. Thisse and Y. Zenou (2000) Wage competition with heterogeneous workers and firms. *Journal of Labor Economics* 18, 453-72.
- [16] Kim, S. (1989) Labor specialization and the extent of the market. *Journal of Political Economy* 97: 692-705.
- [17] MacDonald, D.N. and M.O. Reynolds (1994) Are baseball players paid their marginal product? *Managerial and Decision Economics* 15, 443-457.
- [18] Manning, A. (2003) *Monopsony in Motion. Imperfect Competition in Labor Markets*. Princeton: Princeton University Press.
- [19] Marimon, R. and F. Zilibotti (1999) Unemployment versus mismatch of talents: reconsidering unemployment benefits. *Economic Journal* 109, 266-291.
- [20] Matsuyama, K. (1995) Complementarities and cumulative processes in models of monopolistic competition. *Journal of Economic Literature* 33, 701-729.
- [21] McFadden, D. (1976) The mathematical theory of demand models. In P.R. Stopher and A.H. Meyburg (eds.), *Behavioral Travel-Demand Models*. Toronto, Lexington Books.
- [22] Pigou, A.C. (1924) *The Economics of Welfare*. 2nd ed., London: Macmillan.
- [23] Ransom, M.R. (1993) Seniority and monopsony in the academic labor market. *American Economic Review* 83, 221-233.
- [24] Rosen, S. (2002) Markets and diversity. *American Economic Review* 92, 1-15.
- [25] Salop, S.C. (1979) Monopolistic competition with an outside good. *Bell Journal of Economics* 10, 141-56.

- [26] To, T. (2009) Monopsonistic competition in formal and informal labor markets. In R. Kanbur and J. Svejnar (eds.), *Labor Markets and Economic Development*. U.K., Routledge.