

Filippin, Antonio

**Working Paper**

## Can workers' expectations account for the persistence of discrimination?

IZA Discussion Papers, No. 4490

**Provided in Cooperation with:**

IZA – Institute of Labor Economics

*Suggested Citation:* Filippin, Antonio (2009) : Can workers' expectations account for the persistence of discrimination?, IZA Discussion Papers, No. 4490, Institute for the Study of Labor (IZA), Bonn, <https://nbn-resolving.de/urn:nbn:de:101:1-20091026105>

This Version is available at:

<https://hdl.handle.net/10419/35929>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

IZA DP No. 4490

## Can Workers' Expectations Account for the Persistence of Discrimination?

Antonio Filippin

October 2009

# Can Workers' Expectations Account for the Persistence of Discrimination?

**Antonio Filippin**

*University of Milan  
and IZA*

Discussion Paper No. 4490  
October 2009

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0  
Fax: +49-228-3894-180  
E-mail: [iza@iza.org](mailto: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

### Can Workers' Expectations Account for the Persistence of Discrimination?\*

The paper explains how workers' expectations of being discriminated against can be self-confirming, accounting for the persistence of unequal outcomes in the labour market even beyond the causes that originally generated them. The theoretical framework used is a two-stage game of incomplete information in which one employer promotes only one among two workers after having observed their productivity, which is used as a signal of their ability. Workers who expect to be discriminated against exert a lower effort on average, because of a lower expected return, thereby being promoted less frequently even by unbiased employers. This implies that achievements of minority groups may not improve when the fraction of discriminatory employers actually decreases, and such a mechanism is robust both to trial work periods and to affirmative actions like quotas.

JEL Classification: J71, J15, J24, D82, C79

Keywords: discrimination, workers' expectations, self-confirming beliefs

Corresponding author:

Antonio Filippin  
University of Milan  
Department of Economics  
via Conservatorio 7  
20122 Milano  
Italy  
E-mail: [antonio.filippin@unimi.it](mailto:antonio.filippin@unimi.it)

---

\* A previous version of this paper circulated with the title "Discrimination and Workers' Expectations." I would like to thank participants to the EEA, RES, SYME, AIEL, ASSET conferences, the Brucchi Luchino, the II Gender Economics Workshops (Granada), the 1st Bolzano Summer School in Game Theory and its Applications, as well as seminar participants at the European University Institute, IZA and the universities of Milan, Turin and Cagliari for their valuable suggestions. I thank in particular Pierpaolo Battigalli, Thomas Bauer, Andrea Ichino and Peter Norman. All the remaining errors are mine.

# 1 Introduction

Despite the effort undertaken in fighting discrimination there is a huge empirical literature reporting that there are groups of workers characterized by unequal outcomes even nowadays. Such a persistence is not only hard to tackle in the field, but also difficult to account for from a theoretical point of view. In fact, despite the several contributions to the literature, a widely shared explanation for the long-run persistence of discrimination in the labour market is still not apparent.

The neoclassical theory of discrimination is mostly a demand-side theory. Very few are the contributions where workers' heterogeneity matters and only a paper by Breen and Garcia-Penalosa (2002) studies the possibility that unequal outcomes may persist due to reasons related to workers' expectations. There is instead empirical literature pointing toward an important role played by workers' expectations of being discriminated against. For instance, Filippin (2009b) provides experimental evidence that subjects belonging to populations randomly generated in the lab are likely to show a lower propensity to bid in an all-pay auction after having been discriminated against in the award of the prize. Hoff and Pandey (2004) conducted a field experiment recruiting children belonging to different Indian castes with the task of solving mazes. They show that a large and robust caste gap emerges only when castes are publicly revealed, as long as there is scope for discretion and judgment in rewarding performances.

The goal of this paper is to investigate from a theoretical point of view the role of expectations of being discriminated against in explaining the long-run persistence of unequal outcomes that characterizes some groups of workers. The model is formalized as a two-stage game of incomplete information in which three players participate: two workers, one of which belongs to a minority group, and one employer. The employer promotes one (and only one) of the two workers after having observed their output, which is used as a signal for unobservable ability. Crucially, promotions depend via effort on workers' expectations about the unknown employer's type, which captures the possible disutility of promoting a minority worker, i.e. discriminatory tastes.

What can happen is that unequal outcomes may be supported by minority workers' wrong but self-confirming expectations even though employers do not discriminate against them. The intuition is that minority workers who expect to be discriminated against exert a lower effort on average, because of a lower expected return, inducing a lower percentage of promotions even by unbiased employers. In turn, the outcome is consistent with their beliefs that employers are characterized by discriminatory tastes.<sup>1</sup>

---

<sup>1</sup>Interpreting the choice in the first period as the effort workers exert in acquiring signals, or their job search intensity, and employers' decision as hiring or assigning applicants to different job types, we would obtain that wrong expectations of being discriminated

The model has been designed in order to verify whether workers' expectations stand alone in explaining such a result from a theoretical point of view, but this does not mean that other causes behind unequal outcomes should be regarded as negligible for at least two reasons. First, the rise of discrimination must be explained by other factors, such as discriminatory tastes or statistical discrimination, while workers' expectations can only shed some light on why historically oppressed groups are not likely to forge ahead once the original causes of inequality have been removed. Second, workers' expectations are very likely to interact with statistical discrimination, and the two factors may reinforce each other, making wrong beliefs of both employers and employees self-confirming. Workers' expectations of being discriminated against can also explain why trial work periods are ineffective in fighting statistical discrimination when observing workers' performance is costless.

Although workers' expectations do not account for the rise of discrimination, their importance should not be downplayed because they can be crucial in explaining its persistence, not only from the theoretical point of view, but also as far as the policy interventions are concerned. In fact, unequal outcomes can survive affirmative actions aimed at the demand side of the labour market (e.g. quotas) as long as they do not affect minorities' expectations of being discriminated against.

The structure of the paper is as follows. After clarifying what is meant by discrimination in this paper, Section 2 presents some related literature. The model is explained in Section 3, focusing first on the constituent game, i.e. the game after the players have been matched (Section 3.1), and then on the population game, the matching process and the information structure (Section 3.2), necessary to characterize beliefs about types (Section 3.3) and behavioural rules (Section 3.4). Section 4 analyzes the equilibria and draws some policy implications. Finally, Section 5 concludes.

## 2 Literature Review

Before starting with the presentation of the extensive literature about discrimination, it is useful to clarify what is meant by discrimination in the labour market, since many different and occasionally contradicting definitions have been used. Discrimination has been used implying both different achievements (wages, promotions) for *equally productive* workers, and different achievements for *ex ante equal* workers, i.e. for workers with the same ability and tastes for work. A good compromise, referring in part to Blau, Ferber and Winkler (2002), is to use two different definitions. On the one hand, following the "equal pay for equal work" principle, **direct**

---

against are self-confirming in an equilibrium characterized by segregation across jobs.

**discrimination** can be defined as a different treatment in terms of wages, promotions, or job allocations for *equally productive* workers. On the other hand, a more comprehensive definition seems necessary, too. The reason is that it would be hard to consider as discriminatory an employer who pays or promotes minority workers less (on average) if they are (on average) proportionally less productive. Nevertheless, it would be misleading to disregard that many factors, and direct discrimination can be among the most important ones, may affect workers' behavior. If minority workers are less productive for example because they have changed their behavior reacting to a worse job assignment, the different achievements should not be viewed as equal treatment, even if there is not evidence of direct discrimination. Such a situation is captured by the more comprehensive concept of **cumulative discrimination**, defined as different achievements for *ex ante equal* workers.<sup>2</sup>

Most of the discrimination literature focuses on wages, but the main stylized facts hold in a framework in which workers compete to obtain a promotion, too.<sup>3</sup>

**Discriminatory Tastes.** The starting point of the economic analysis of discrimination in the labour markets is the article "The Economics of Discrimination" by Becker (1957). In Becker's model, the existence of *direct discrimination* between workers of different groups, which are perfect substitutes in the production function, is based on the discriminatory preferences of employers, coworkers or customers. Hence, discrimination is caused by fundamentals (discriminatory tastes), while beliefs do not play any role because there is no uncertainty. Within this framework, members of the discriminated group must receive a lower wage in order to be accepted as employees, coworkers or sales. Becker's approach can explain the rise of any type of direct discrimination (based on sex, race, religion, etc.). The major problem lies in its long run implications: if markets are competitive and there is heterogeneity of discriminatory tastes, only the less discriminatory employers (or the non-discriminatory ones if present) should survive because

---

<sup>2</sup>Another distinction that deserves to be mentioned is that between *group* and *individual* discrimination. The latter happens when an individual is judged also on the basis of group membership rather than upon his own characteristics only. Individual discrimination is a characteristic of all the models of incomplete information, it concerns both the majority and the minority group, and it does not imply group discrimination. Henceforth, even though not specified, discrimination always refers to group discrimination.

<sup>3</sup>Theories have been selected in order to facilitate contrast and comparison with the model of workers' expectations, mainly focusing on the theoretical aspects of some competitive neoclassical models and institutional theories. Also the relative weights assigned to various aspects of such theories reflect primarily the necessity of the subsequent presentation rather than some sort of consensus. Another reason for these choices is that many detailed surveys are already available (see Blau, Ferber and Winkler (2006) and Cain (1986) among others).

discrimination is costly. Alternatively, we should observe complete segregation. However, both predictions are contradicted by empirical evidence.

**Statistical Discrimination.** In the statistical discrimination models, group membership is assumed to convey information regarding individual characteristics about which incomplete information is assumed. Several models have been developed within this strand of literature using different devices in order to explain the long run persistence of observed discrimination.<sup>4</sup> Common to these models is the fact that, differently from Becker's one, fundamentals are not relevant. Employer's beliefs about the existence of different characteristics between groups of workers that are instead *ex ante* identical turn out to be correct in equilibrium. Why are these wrong beliefs self-confirming? The mechanism is the following: a worker's a priori unobservable variable (e.g. effort) is endogenously affected by employer's beliefs (e.g. via lower wages, or via worse job assignments), leading to a suboptimal choice of effort that determines an outcome consistent with the beliefs of the employer. The conclusion is that there is cumulative but not direct discrimination, because workers are *ex ante* equal but display a different productivity in equilibrium.

Statistical discrimination models have been criticized by Aigner and Cain (1977) and Cain (1986), on the ground that "employer's uncertainty about the productivity of workers may be inexpensively reduced by observing the workers' on-the-job performance." This argument does not apply when observing workers' productivity implies a sunk cost, as in De la Rica, Dolado and Penalosa (2009). But even when observing workers' performance is costless what determines minority workers' behavior in the trial work period? In other words, is it optimal for them to increase their effort to be assigned to the good job? The answer cannot be found within the statistical discrimination literature, because it is necessary to analyze also the role of workers' expectations. In section 4, it emerges that trial work periods are not an effective policy device to break down unequal outcomes, as long as minority workers believe they are discriminated against.

**The Human Capital Theory.** Another strand of the literature, started by Mincer and Polacheck (1974) analyzes the effects of voluntary choices of investment in human capital from a gender perspective. According to this theory, women decide to invest less than men because they expect a lower lifetime return on human capital due to a shorter and more discontinuous presence in the labour force. As a consequence, they receive less on-the-job training and/or are assigned to less rewarding jobs. Such behavior can be

---

<sup>4</sup>The seminal contribution in the statistical discrimination literature has been proposed by Arrow (1973). Other examples of statistical discrimination models can be found in Phelps (1972), who concentrates on the effect of an imperfect predictor of the true productivity of a worker, and Spence (1973), in his pioneer work about signaling. More recently, Moro and Norman (2004) analyze statistical discrimination using a general equilibrium approach.



ascribed to the traditional division of work within the family (Becker, 1985). In this way, wage differentials, worse career path, and sex segregation are explained by voluntary choices. If this is the case, the different achievements could not be classified as discrimination, given that workers neither equally productive in equilibrium nor *ex ante* equal.

**Feedback Effects.** Some economists have heavily criticized the human capital approach because the seemingly “voluntary” decision could instead be induced by discrimination, entering the definition of cumulative discrimination. For instance, Blau and Jusenius (1976), reverse the causality link with respect to Mincer and Polacheck (1974): women, because of experiences of sex discrimination, e.g. lower wages, respond with career interruptions and specialization in household production, i.e. investing less in human capital. A large number of the so called “institutional” contributions may also fall into this category, starting from Myrdal (1944), who takes the behavior of workers explicitly into account: “Negro worker often feels that his fate depends less on his individual efforts than on what white people believe about Negroes in general.” Ferber and Lowry (1976) also follows along the line of such a vicious circle. The explicit analysis of workers’ expectations in this paper can be seen as a way to formalize such feedback effects, that are usually stated in a qualitative way.

**Workers’ expectations.** The only paper focusing explicitly on the supply side of the labour market is that of Breen and Garcia-Penalosa (2002), who explain the observed persistence of gender segregation using a Bayesian learning approach. Workers, due to imperfect information, do not know and try to learn how much the probability of success in various occupations is affected by effort or by predetermined individual characteristics (such as gender). The “prior” of a man (woman) is the belief received by his father (her mother), while the posterior is the belief updated according to his (her) experience and transmitted to his son (her daughter). Different preferences between men and women at some point in the past caused different learning paths and different beliefs. This is a sufficient condition to observe lasting unequal outcomes in equilibrium for the two groups, even once preferences become equal, meaning that past circumstances continue to exert an influence and that expectations can be self-fulfilling. What differs with respect to the model presented in Section 3 is the information structure. Agents learn from their parents only, but not from observable aggregate outcomes. Moreover, only agents choosing a “high” profile of education and effort are able to learn from their experience and transmit updated beliefs to their children, while for the “low” profile the learning process stops. The information structure of the model prevents agents from learning that differences in fundamentals have disappeared, so that beliefs are still a function of differences in workers’ fundamentals. Section 4 will show that workers’ expectations can explain observed unequal outcomes within a static framework in which such an assumption is relaxed allowing agents to access aggregate information.

**Asymmetric Tournaments.** A tournament is symmetric when outcomes are invariant to the permutation of the contestants. On the other hand, asymmetric contests are defined “uneven” when agents are different, and “unfair” when contestants are identical but the rules favour one of them. The literature on tournaments, started by Lazear and Rosen (1981), is not directly related to discrimination. Nevertheless, asymmetric tournaments as described by O’Keeffe et al. (1984) provide a useful benchmark for the analysis of the effects of discrimination on promotions, and not surprisingly the two models provide similar prediction in some cases, e.g. that discriminatory tastes affect the incentives of majority and minority workers in the same way. Within the literature on uneven tournaments, it is incidentally mentioned that unequal outcomes may arise between equal groups of workers.<sup>5</sup> However, the underlying mechanism has not been formalized and no role is explicitly played by expectations. Hence, the result presented in Section 4 that effort differs across otherwise identical workers because of their different beliefs may also be interpreted as a formal justification for the existence of uneven tournaments between *ex ante* identical participants.

### 3 The Model

The model is formalized as a two-stage game of incomplete information where populations of workers and employers are engaged. The two populations of workers differ because of an observable characteristic (race, gender, etc.) that, without affecting their productivity, distinguishes the so-called majority workers, identified by subscript  $A$ , from the so-called minority workers (subscript  $B$ ). Employers are denoted by subscript  $F$ .

The following section focuses on the constituent game, i.e. on what happens after the players have been drawn from their populations and matched. The population game, the matching process and the information structure, necessary to characterize beliefs, are described in Section 3.2.

#### 3.1 The constituent game

In every constituent game two workers, one of which is a “minority” worker, and one employer are drawn from their populations and play a two-stage game. In the first period both workers choose a level of effort  $(e_A^1, e_B^1)$ , then the employer promotes one (and only one) of them after having observed their productivity  $(\pi_A^1, \pi_B^1)$ . After the employer’s decision, workers choose a level of effort for the second period  $(e_A^2, e_B^2)$ .

The labour market is assumed to be competitive, hence the productivity is entirely paid to workers ( $w = \pi$ ). This assumption makes the game equivalent to the reduced form of a more general game where workers’ output

---

<sup>5</sup>See Schotter and Weigelt (1992).

is observed and verifiable and employers compete on enforceable piece-rate contracts. Workers are free to move, but in equilibrium  $w = \pi$  and nobody moves.

### 3.1.1 Incomplete Information

The game is characterized by incomplete information, because every player knows his type only:

$\theta_A \in [1, 2]$  and  $\theta_B \in [1, 2]$  summarize the ability of the majority and of the minority worker, respectively, that are identical on average.

$\theta_F \in \{0, d\}$  represents employer's tastes for discrimination. There are only two types of employer. If  $\theta_F = 0$  (henceforth:  $\theta_0$ ) the employer is unbiased, meaning that the observable characteristic that distinguishes the two workers does not affect his payoffs. On the other hand if  $\theta_F = d$  (henceforth:  $\theta_d$ ) the employer suffers a disutility when the minority worker is promoted.  $\theta_d$  is assumed to be sufficiently large that promoting worker A is always the optimal choice regardless of workers' productivity. Since the labour market is competitive, discrimination in this model can only take place denying a promotion to a worker that deserves it.

Summarizing, a minority worker knows his own ability  $\theta_B$ , while the ability of the majority worker  $\theta_A$  and the tastes for discrimination of the employer  $\theta_F$  are unknown. The same holds *mutatis mutandis* for players  $B$  and  $F$ .

### 3.1.2 Payoffs

The structure of the utility function is the same for majority ( $A$ ) and minority ( $B$ ) workers:

$$U_P = w_P^1 - \frac{(e_P^1)^2}{K^1} + w_P^2 - \frac{(e_P^2)^2}{K^2}, \quad (1)$$

where:

$U_P$  is the utility of a worker belonging to population  $P = \{A, B\}$ ;

$e_P^t$  and  $w_P^t$  are effort and wage in period  $t$ , respectively;

$K^1$  and  $K^2$  are two constants that weight the disutility of effort in the two periods. Although the two constants could take any value, including one, I assume that  $K^1 = 2$  and  $K^2 = 4$  in order to obtain simpler equations for the optimal choice of effort.

As already mentioned, productivity is entirely paid to workers because the labour market is assumed to be competitive. What has to be explained yet is how ability and effort interact in generating the output  $\pi$ .

I assume that productivity always coincides with effort, except for the promoted worker in whose case ability also matters.<sup>6</sup>

$$\pi_P^1 = e_P^1; \quad \pi_P^2(\alpha_P) = \sqrt{\theta_P} e_P^2; \quad \pi_P^2(\alpha_{-P}) = e_P^2, \quad (2)$$

where  $\alpha_P$  means that the worker belonging to the population  $P = \{A, B\}$  is promoted, while  $\alpha_{-P}$  means that the opponent is promoted. Such a relation between effort, ability and productivity is the most parsimonious version of the model that delivers some nice features. The first is that promotions are desirable, because as long as  $\theta > 1$  the wage turns out to be higher when the worker is promoted *ceteris paribus*. The second is that ability matters more for high-skill jobs.

Substituting (2),  $K^1 = 2$  and  $K^2 = 4$  into (1) we get:

$$U_P = e_P^1 - \frac{(e_P^1)^2}{2} + I(\alpha_P) \left[ \sqrt{\theta_P} e_P^2 - \frac{(e_P^2)^2}{4} \right] + I(\alpha_{-P}) \left[ e_P^2 - \frac{(e_P^2)^2}{4} \right], \quad (3)$$

where  $I(\cdot)$  is an indicator function that takes the value 1 when its argument is true. This specification of the utility function is equivalent to that used by O'Keefe et al. (1984) dealing with tournaments. The only difference is that here the role of the prize is played by the higher wage attached to the job of the promoted worker.<sup>7</sup>

As far as the employer is concerned, the utility function contains both monetary earnings and a parameter summarizing the possible disutility associated to the promotion of worker  $B$ , who is the only one that faces the risk of being discriminated against.

Employer's utility function is:

$$U^F = (m - 1)(\pi_A^1 + \pi_B^1 + \pi_A^2 + \pi_B^2) - I(\alpha_B)\theta_F$$

where  $m > 1$  is a known and constant mark up on workers' productivity due to the entrepreneurial activity. The term  $I(\alpha_B)\theta_F$  represents the disutility associated to a promotion of a minority worker. When  $\theta_F = 0$  the observable characteristic that distinguishes the workers does not matter and earnings are the only thing that the employer considers. On the other hand, when  $\theta_F = d$  the employer is characterized by discriminatory tastes and is willing to give up monetary earnings in order to avoid the minority worker being employed in the high-skill position. Therefore, in order to maximize earnings, the employer needs to maximize worker's productivity. In other words, the employer has an incentive to promote the more productive worker. Note that

<sup>6</sup>The square root is used to simplify as much as possible the expression for the optimal level of effort.

<sup>7</sup>Such a specification is fairly general, although it cannot account for catching-up effects that would emerge if minorities were characterized by a backward sloping labour supply.

employer's earnings are not a simple return on capital, but a compensation for an activity actually undertaken, namely that of organizing production factors. In this model positive earnings are not a necessary condition to support discrimination. In contrast, they give the incentive to promote the more productive worker regardless of her population.<sup>8</sup>

### 3.1.3 Strategies

Workers move twice, the second time after the decision about promotion has been taken by the employer, choosing effort simultaneously. The strategy  $s$  of a worker is therefore a triple containing an effort level for the first period, and two effort levels for the second period, one if promoted, another if not promoted.

Applying backward induction is straightforward to compute that in the second period the optimal effort is higher when the worker is promoted:

$$e_P^{2*}|\alpha_P = 2\sqrt{\theta_P} \geq e_P^{2*}|\alpha_{-P} = 2. \quad (4)$$

The choice in the first period is more interesting to analyze. Considering the instantaneous utility the optimal effort would be  $e_P^{1*} = 1$ . However, the worker is willing to accept a loss of utility in the first period in order to increase her chances of being promoted, because if assigned to the high-skill job he experiences a gain of utility (increasing in ability). Therefore, the promotion acts as incentive device, and effort in the first period can be used as a signal of workers' ability: the higher the ability, the higher the loss of utility accepted in the first period via a suboptimally higher effort.<sup>9</sup> Exploiting equations (4) workers' lifetime expected utility can be written:

$$EU_P(e_P^1) = e_P^1 - \frac{(e_P^1)^2}{2} + 1 + Pr[\alpha_P(e_P^1)](\theta_P - 1), \quad (5)$$

where:

$\theta_P - 1 > 0$  is the gain of utility when promoted, and

$Pr[\alpha_P(e_P^1)]$  is the probability of being promoted, which also depends on the level of effort exerted in the first period.

Note that when there is no chance of being promoted the last term of equation (5) vanishes and the optimal choice becomes  $e_P^{1*} = 1$ . In other

---

<sup>8</sup>It is also possible to interpret  $F$  as a supervisor rather than an employer. Instead of profits, the supervisor maximizes a bonus which is a fraction of the overall productivity of the workers. Nothing changes in practice, because also the supervisor has the incentive to promote the more productive worker in order to maximize her bonus.

<sup>9</sup>To refine the set of equilibria it is useful to assume that each worker thinks that the probability of being promoted cannot be decreasing in effort *ceteris paribus*. In principle, nothing prevents players from using a suboptimally lower (instead of higher) level of effort as a signaling device. However, such equilibria do not provide additional insights and are therefore disregarded.

words, in a world without promotions, as well as when all the employers are discriminatory, only the instantaneous utility in the first period matters. Taking first order conditions of (5) we find the optimal choice of effort in the first period, which depends on its impact on the probability of being promoted, weighted according to worker's ability:

$$e_P^{1*} = 1 + \frac{\partial Pr[\alpha_P(e_P^1)]}{\partial e_P^1}(\theta_P - 1). \quad (6)$$

Note also that, *ceteris paribus*, the higher  $\theta$ , the higher the optimal effort in the first period, which, as already mentioned, can be interpreted as a signal of ability.

The set of feasible actions of the employer is simply  $A_F = \{\alpha_A, \alpha_B\}$ , where  $\alpha_A$  stands for “promotes worker  $A$ ” and  $\alpha_B$  stands for “promotes worker  $B$ .” The set of strategies  $S_F$  can be summarized by means of a correspondence that specifies a promotion decision for every possible combination of productivity (effort) levels observed in the first period.

Employers of type  $\theta_F = d$  are characterized by a disutility of promoting a minority worker that is assumed for the sake of simplicity to be so high that they always promote worker  $A$ :

$$\alpha_A = BR^{\theta_d}(\pi^1)$$

Employers of type  $\theta_F = 0$  are not affected by the observable characteristic that distinguishes workers  $A$  from workers  $B$ , and therefore they do not suffer a disutility promoting a minority worker. It follows that the best reply depends on the comparison of workers' expected productivity in the second period. Note that only workers' difference in productivity *after* the promotion affects employer's decision, while the difference in the first period is sunk. Therefore, at the margin only benefits from the promotion of a minority worker (i.e. the difference in productivity in the second period) are compared with the associated cost  $\theta_F$ . Defining  $\mu_F(\theta_A|\pi_A^1)$  the updated beliefs the employer holds about the ability of worker  $A$  having observed productivity  $\pi_A^1$ , the best reply is:

$$\alpha_A \in BR^{\theta_0}(\pi^1) \quad \text{if} \quad \int \pi_A^2 d\mu_F(\theta_A|\pi_A^1) > \int \pi_B^2 d\mu_F(\theta_B|\pi_B^1). \quad (7)$$

This simply means that after having inferred workers' ability from the signal provided by the observed productivity, promoting a majority worker is the best reply whenever her expected productivity is higher. Of course, the same holds as far as the minority worker is concerned when equation (7) holds with a reversed inequality sign. Since productivity is increasing in ability in

both periods (see equations 4 and 6) the best reply of a non-discriminatory employer is to implement a fair tournament and to promote the worker displaying the higher productivity in the first period.<sup>10</sup>

Employers' behavioural rules allow to better characterize workers' optimal choice of effort. In fact, the probability of being promoted can be rewritten:

$$\begin{aligned} Pr(\alpha_B) &= Pr(\theta_0)Pr(e_B^1 > e_A^1), \\ Pr(\alpha_A) &= Pr(\theta_0)Pr(e_A^1 > e_B^1) + 1 - Pr(\theta_0), \end{aligned}$$

i.e. as a function of the percentage of non-discriminatory employers, as well as of the fraction of contestants against whom the worker would win the tournament. The latter is simply the cumulative distribution of effort in the other population evaluated at one's own level of effort. The incentive to exert an effort higher than one, however, depends on the *change* of the probability of being promoted and not on its *level* (see equation 6). Therefore, the optimal choice of effort will contain the density of effort in the population of opponents evaluated at the level of effort chosen by the worker:

$$e_P^{1*} = 1 + Pr(\theta_0) f_{e_{-P}^1}(e_P^{1*}) (\theta_P - 1). \quad (8)$$

Note that the incentives to exert effort are the same for both populations even when there are discriminatory employers, in line with the literature on asymmetric tournaments. This is intuitive, because the assumption that discriminatory employers always promote worker *A* implies that incentives are proportional to the percentage of non-discriminatory employers for both workers. In the limit situation where there are only discriminatory employers, promotion stops to be an incentive device for both workers, because *A* is sure to be promoted, while *B* has no chance.

To complete the description of the constituent game, also players' beliefs need to be specified. In order to do this, however, it is necessary to describe how players are matched and the information they can access.

### 3.2 The Population Game

The constituent game described in section 3.1 is inserted in a wider game, called population game, from which the information structure that allows to define players' beliefs arises.

There are three populations, one of employers and two of workers. Like in the constituent game the populations of workers differ because of an

---

<sup>10</sup>In the next section it is explained that this is true only if the employer does not figure out that workers hold different expectations about the fraction of discriminatory employers.

observable characteristic (gender, race, etc.) that does not affect their productivity. The distribution of types within the two populations of workers is identical but this is not common knowledge. This assumption rules out the possibility that unbalanced promotions across populations arise because of a different average ability.

### 3.2.1 Matching

Each of the three populations is composed by a continuum of players. At every stage each player of population  $A$  is randomly matched with one player from population  $B$  and one player from population  $F$ .

### 3.2.2 Information Structure

Players infer types and strategies of their opponents using available information. Besides observing the history  $\omega = (\pi_A^1, \pi_B^1, \alpha_P, \pi_A^2, \pi_B^2) \in \Omega$  of his constituent game, every player is assumed to observe the distribution of promotions between populations  $\hat{\alpha} \in \Delta(A_F)$  arising from the whole population game, i.e. also from the constituent games in which the player is not directly involved. At first sight individual information seems to be more informative, since it allows players to compute the probability of being promoted conditional on workers' output, while aggregate information does not. However, if workers participate in the labour market for a small number of rounds, as it is plausible to assume, individual information becomes negligible. Learning from the same aggregate information is what makes the information structure of this model more general than that in Breen and Garcia-Penalosa (2002) already described in Section 2. However, it is crucial that only the distribution of promotions is observed while the distribution of output is not, otherwise minorities should rationalize when unbalanced promotions are not due to discriminatory tastes.

## 3.3 Exogenous Beliefs

The Common Prior assumption implies that in a game with three players every pair of players must hold common beliefs about the type of the third. Although debated from a theoretical point of view,<sup>11</sup> this assumption is usually made in Bayesian games. Empirical evidence, however, suggests that this is not the case dealing with workers' expectations. In a dataset collected by Filippin and Ichino (2005) among 2nd year Bocconi students, male and female share similar expectations about the magnitude of the gender wage gap after graduation, but the reasons differ considerably. While discriminatory tastes is the explanation chosen by one third of the females and 20 percent of males, 10 percent more males attributes the gap to actual

---

<sup>11</sup>See the critique of Gul (1998) and the reply of Aumann (1998).



gender differences. I internalize this finding allowing majority and minority workers to hold different beliefs about the type of the employer, thereby relaxing the Common Prior assumption.

Although the setting of this game is static, a thought experiment that provides a dynamic intuition about how beliefs are formed is useful. Beliefs at time  $t$  can be thought as a function of the information arising from the population game in  $t - 1$ . For instance, workers can infer the fraction of discriminatory employers using the distribution of promotions  $\hat{\alpha}_{t-1}$ . Hence, workers  $A$  and  $B$  can interpret a given distribution of promotions across populations assigning a different importance to the role played by employers' propensity to discriminate against the minority.

Besides differing across players, beliefs are not assumed to be common knowledge; hence, players must hold second order beliefs about opponents' expectations.<sup>12</sup>

Exogenous beliefs of a player are a probability measure over the unknown component of the belief-type set  $M \times \Theta$ . Given that every player is supposed to know his type and his beliefs, the unknown component of  $M \times \Theta$  turns out to be the set of belief-type profiles of all the other players, both the opponents and the other players of his own population.<sup>13</sup> Exogenous beliefs of a worker of type  $\theta$  of population  $P$  are therefore defined

$$\mu_{\theta_P} \in \Delta(M \times \Theta).$$

What follows is a list of assumptions that simplifies the information structure used by workers' to form their expectations:<sup>14</sup>

1. *Belief-type profiles of opponents are not correlated.* In other words, opponents do not correlate their beliefs:

$$\mu_A(\mu_B, \theta_B, \mu_F, \theta_F) = \mu_A(\mu_B, \theta_B) \mu_A(\mu_F, \theta_F).$$

2. *Beliefs are type-independent,* i.e. players belonging to the same population share the same beliefs:

$$\mu_{\theta_A} = \mu_A.$$

This makes unnecessary to specify *beliefs* in the profile of the opponents, and now I simply refer to *types*.

---

<sup>12</sup>The game as presented below could equivalently be described defining epistemic types *a la* Harsanyi (1967-68), i.e. specifying a (common knowledge) universal type-space that summarizes players' uncertainty and from which an infinite hierarchy of beliefs can be derived.

<sup>13</sup>In the constituent game each player can disregard the beliefs of the other players of his population. However, such beliefs become relevant when trying to rationalize  $\hat{\alpha}$ , the observed distribution of promotions between populations, that workers can use to infer the fraction of discriminatory employers.

<sup>14</sup>Beliefs are explicitly shown for a player belonging to population A, but the same holds for populations B and F *mutatis mutandis*.

The following is the full specification of players beliefs that, although admittedly tedious, is necessary to characterize the equilibria in Section 4 below.

### 3.3.1 Minority Workers

Expectations of minority workers are assumed to be correct as far as the distribution of ability is concerned, while beliefs concerning the fraction of discriminatory employers can be wrong. They attribute the whole gap in the promotion rate to the discriminatory tastes of the employers, and they believe that their opponents share the same belief as them. In more detail:

1. Workers  $B$  infer the fraction of discriminatory employers from the distribution of promotions across populations:

$$\mu_B(\theta_d) = 2(0.5 - \hat{\alpha}_B). \quad (9)$$

Only when promotions are balanced across populations ( $\hat{\alpha}_B = 0.5$ ) they think that all the employers are unbiased ( $\mu_B(\theta_d) = 0$ ).

2. Minority workers correctly believe that the two populations of workers are characterized by the same uniform distribution of ability:

$$\mu_B(\theta_A) \sim U[1, 2]; \quad \mu_B(\theta_B) \sim U[1, 2].$$

3. Minority workers cannot be so naive to infer the fraction of discriminatory employers from the promotion rate while at the same time believing that the other populations correctly figure out the true fraction of discriminatory employers. Therefore, they are assumed to think that the same belief rule applies to the other players as well:

$$\mu_B [\mu_A(\theta_d)] = \mu_B [\mu_F(\theta_d)] = \mu_B(\theta_d).^{15}$$

Similarly, they believe that the opponents also think that the two populations are equal.

### 3.3.2 Employers and Majority Workers

Expectations of employers and majority workers concerning the distribution of opponents' type are assumed to be correct. They also believe that their opponents share the same belief as theirs. In particular:

---

<sup>15</sup>  $\mu_B [\mu_A(\theta_d)]$  represent second order beliefs of worker  $B$  over first order beliefs of worker  $A$  about the type of the employer. This is indeed an abuse of notation used for the sake of simplicity. Second order beliefs are defined on both the uncertainty domain and on opponents' first order beliefs, and the correct corresponding notation should therefore be:  $\mu_B^2 \in \Delta [M \times \Theta \times \Delta(M \times \Theta)]$ , see Siniscalchi (2007).

1. Both employers and majority workers hold correct beliefs about the distribution of types in all the three populations:

$$\begin{aligned} \mu_A(\theta_A) \sim U[1, 2] \quad \mu_A(\theta_B) \sim U[1, 2] \quad \mu_A(\theta_d) &= Pr(\theta_d); \\ \mu_F(\theta_A) \sim U[1, 2] \quad \mu_F(\theta_B) \sim U[1, 2] \quad \mu_F(\theta_d) &= Pr(\theta_d). \end{aligned}$$

The assumption that the employer thinks that the two populations of workers are identical rules out the possibility that unequal outcomes are supported by statistical discrimination in this version of the model.

2. Both employers and majority workers think that also all the other players share the same beliefs as theirs. The assumption that employers think that minority workers have correct beliefs about the fraction of discriminatory employers is crucial. In fact, if employers correctly figure out minority workers' beliefs, i.e.  $\mu_F[\mu_B(\theta_d)] = \mu_B(\theta_d) \geq Pr(\theta_d)$ , they would manage to invert minority workers' choice function (Equation 8 above) and correctly determine minority workers' type. This would trivially imply that employers reach complete information about all workers' type at the end of the first period, while, in line with an old saying, some uncertainty is necessary for a horse race to take place.

Note that this is simply the most parsimonious framework that delivers the punchline of the paper, but there are other settings of beliefs that would make sense and that would probably be even more plausible. For instance, beliefs might differ within populations, and wrong expectations could characterize only a subgroup of the minority workers (e.g. those characterized by lower ability).

### 3.4 Endogenous Beliefs

Players fully exploit their knowledge to derive the behavioural rules of each type of the opponents. This means that once a type is identified, expectations about his strategy are derived relying upon the knowledge of the rules of the game as well as common certainty of rationality. For instance, majority workers infer that minority workers choose their effort in the first period according to equation 8. However, substituting what they ignore with the corresponding expectations is not enough to fully characterize the behavioural rule of the opponents. In fact the optimal choice of worker *A* also depends on the strategies of workers *B*, which in turn depend on what workers *A* choose. Further assumptions about (beliefs over) behavioural rules are therefore necessary to select among all the possible equilibria of the model presented so far.

1. Workers internalize that non-discriminatory employers implement a fair tournament.

2. Opponents do not correlate their play in such a way that:

$$\mu_A(\mu_B, s_B, \mu_F, s_F) = \mu_A(\mu_B, s_B)\mu_A(\mu_F, s_F).$$

3. Workers believe that the density of effort of their opponents is uniformly distributed:<sup>16</sup>

$$\mu_A[e_B] = U[1, \mu_A(\max(e_B))]; \quad \mu_B[e_A] = U[1, \mu_B(\max(e_A))].$$

4. In line with the assumption already made concerning opponents' type, minority workers believe that the two populations are identical, and therefore that they behave in the same way:

$$\mu_B(e_A) = u_B. \tag{10}$$

5. Majority workers and employers rationalize possible gaps in the promotion rates across populations adjusting their beliefs about minority workers' behavioural rule. In more detail, if relatively less minority workers are promoted, majority workers and employers infer that the density of effort of minority worker must be thicker (and hence the upper bound of the distribution lower) than that of majority workers:

$$\mu_A(e_B) = \mu_F(e_B) = \frac{1}{\sqrt{2\hat{\alpha}_B}}. \tag{11}$$

Moreover, employers deduct that the density of effort of majority workers must be correspondingly thinner:

$$\mu_F(e_A) = \sqrt{2\hat{\alpha}_B}. \tag{12}$$

## 4 Analysis of the Equilibria

In this section two solutions of the model are presented. The first (see Proposition 1) shows that if the expectations of all players are correct the promotion rate of minority workers is in line with the fundamentals, and in the limit case in which all the employers are unbiased promotions are balanced across populations. The second (see Proposition 2) shows that unequal outcomes are more pronounced when minority workers' expectations are wrong *ceteris paribus*. In this case unequal outcomes can survive the disappearance of discriminatory tastes.

---

<sup>16</sup>To simplify the notation I will refer to the true densities simply as  $u_A$  and  $u_B$ . Moreover, I omit effort in the second period since it is uniquely determined via backward induction. Consequently, I also omit the time superscript because beliefs always refer to effort in the first period.

Two different concepts are necessary to analyze the equilibria of the model: the Perfect Bayesian Equilibrium (henceforth: PBE) and the Self-Confirming Equilibrium (henceforth: SCE). The two equilibrium concepts share the feature that players maximize their utility given beliefs, updated whenever possible, about opponents' profile. The difference is that in a PBE players are sophisticated and hold correct conjectures about the relationship between opponents' types and choices, i.e. about their behavioural rules. In the commonly applied subcase when beliefs satisfy the Common Prior assumption, beliefs about opponents' types are correct as well. On the other hand, when the Common Prior assumption is violated, exogenous beliefs may be contradicted by the evidence, thereby failing of representing a fixed point of a learning process. In contrast, in a SCE usually players are less sophisticated, because even knowledge of rationality is sometimes not assumed. Beliefs may differ from the true distribution of opponents' belief-type-strategy profiles, but they are required to be consistent with the evidence.<sup>17</sup>

Correctness of beliefs is therefore a crucial issue in this game. Beliefs are *correct* whenever, for every player of each population, the subjective probability distribution over opponents' belief-type-strategy profiles coincides with the objective one:

$$\mu(\theta, \mu, s) = Pr(\theta, \mu, s).$$

This is what happens in Proposition 1 below.

As already mentioned, players can interpret in different ways the same information about aggregate outcomes. But if players hold different beliefs, some of them must be wrong. Asymptotic empiricism, however, requires that in equilibrium all the belief rules must generate subjective distributions of observables that coincide with the objective one, thereby representing a fixed point of a learning process. In other words, beliefs must be *consistent with the evidence*, which means that, for every player, the subjective probability of observing a particular distribution of promotions  $\hat{\alpha}$  coincides with the actual frequency. The subjective probability is obtained summing up the probabilities attached to every combination of opponents' belief-type-strategy profiles that leads to the same observables  $\hat{\alpha}$ . Beliefs can therefore be consistent with the evidence even though incorrect, and this is what happens when unequal outcomes persist due to minority workers' wrong expectations. Despite the Common Prior is violated neither beliefs of worker *A* nor beliefs of worker *B*, though different, are contradicted by the evidence.

---

<sup>17</sup>For a thorough exposition of the characteristics of SCE the reader is referred to Battigalli (1987) and Fudenberg and Levine (1993 and 1998). The generalization of the SCE to the case of aggregate outcomes is described in Filippin (2009a). See Battigalli and Guaitoli (1997) and Dekel, Fudenberg and Levine (2004) for a formal discussion of the relation between the Common Prior assumption, PBE and SCE in games of incomplete information.

## 4.1 The Equilibria of the Game

This section analyzes the equilibria of the game focusing on the role of workers' expectations. Considering the assumptions made so far, only one possible difference between workers  $A$  and workers  $B$  survives in the model: their expectations about employers' type. In particular, beliefs of workers  $B$  about employers' type may be correct  $\mu_B(\theta_0) = \Pr(\theta_0)$  or wrong  $\mu_B(\theta_0) < \Pr(\theta_0)$ , where  $Pr(\theta_0)$  is the actual percentage of non-discriminatory employers. Proposition 1 and Proposition 2 contrast what happens in these two different situations, everything else being equal.

**Proposition 1** *When expectations of workers  $B$  about employers type are correct, there exists a unique Self-Confirming and Perfect Bayesian Equilibrium in which:*

- 1) *in the first period minority workers choose the same action of the corresponding type of population  $A$ .*
- 2) *the percentage of minority workers promoted is equal to  $0.5 \Pr(\theta_0)$ .*

When expectations of minority workers about employers type are correct  $\mu_B(\theta_0) = \Pr(\theta_0)$  the optimal choice of minority workers in the first period becomes:

$$e_B^{1*} = 1 + Pr(\theta_0)\mu_B(e_A)(\theta_B - 1),$$

which, exploiting equation (10) as well as that  $max(\theta_B) = 2$ , leads to:

$$\begin{aligned} \mu_B(e_A) &= u_B = \frac{1}{\sqrt{\Pr(\theta_0)}} \quad \text{and therefore} \\ e_B^{1*} &= 1 + \sqrt{\Pr(\theta_0)}(\theta_B - 1). \end{aligned} \quad (13)$$

According to (11) the optimal choice of majority worker turns out to be:

$$e_A^{1*} = 1 + \frac{\Pr(\theta_0)}{\sqrt{2\hat{\alpha}_B}}(\theta_A - 1); \quad (14)$$

$$\text{and therefore } u_A = \frac{\sqrt{2\hat{\alpha}_B}}{\Pr(\theta_0)}.$$

Since the minimum effort in both populations is equal to one, the actual fraction of minority workers promoted will be half of those whose distribution of effort overlaps that of majority workers provided that they face an unbiased employer. On the other hand, all the majority workers exerting an effort higher than the maximum of the minority workers, as well as those who are matched with discriminatory employers will be promoted. Therefore:

$$\begin{aligned} \hat{\alpha}_A &= \Pr(\theta_d) + \Pr(\theta_0) \frac{max(e_A^1) - max(e_B^1)}{max(e_A^1) - 1} + 0.5 \Pr(\theta_0) \frac{max(e_B^1) - 1}{max(e_A^1) - 1} \\ \hat{\alpha}_B &= 0.5 \Pr(\theta_0) \frac{max(e_B^1) - 1}{max(e_A^1) - 1}. \end{aligned} \quad (15)$$

Combining (13), (14) and (15) we get that  $e_A^{1*} = e_B^{1*}$  and that:

$$\hat{\alpha}_B = 0.5 \Pr(\theta_0).$$

All players maximize their utility given their beliefs, which are not contradicted by the evidence. The fraction of minority workers promoted depends on the actual fraction of discriminatory employers. If discriminatory tastes disappear, i.e.  $\Pr(\theta_0) = 1$ , promotions are balanced across populations. Notice that the optimal choice of effort is the same for both workers, because biased employers affects their incentives in the same way. The higher the fraction of discriminatory employers, the lower the importance of effort in determining the outcome of the tournament

What happens instead if workers  $B$  overestimate the percentage of discriminatory employers? Assuming that  $\mu_B(\theta_0) < \Pr(\theta_0)$  while  $\mu_A(\theta_0) = \Pr(\theta_0)$  violates the Common Prior assumption, so that beliefs about employers' type must be incorrect for some players and may also be contradicted by the evidence. However, Proposition 2 states that an equilibrium exists satisfying the requirements of both the PBE and the SCE, in such a way that beliefs, although wrong, are consistent with the evidence.

**Proposition 2** *When minority workers overestimate the percentage of discriminatory employers, there exists a unique Self-Confirming and Perfect Bayesian Equilibrium in which:*

- 1) *in the first period minority workers choose a level of effort lower than that of the corresponding type of population A.*
- 2) *the percentage of minority workers promoted is equal to  $0.5\mu_B(\theta_0)$ .*

If  $\mu_B(\theta_0) < \Pr(\theta_0)$  the optimal choice of minority workers in the first period becomes:

$$e_B^{1*} = 1 + \mu_B(\theta_0)\mu_B(e_A)(\theta_B - 1), \quad (16)$$

which, exploiting equation (10) as well as that  $\max(\theta_B) = 2$ , leads to:

$$\begin{aligned} \mu_B(e_A) &= u_B = \frac{1}{\sqrt{\mu_B\theta_0}} \quad \text{and therefore} \\ e_B^{1*} &= 1 + \sqrt{\mu_B(\theta_0)}(\theta_B - 1). \end{aligned} \quad (17)$$

The optimal choice of majority worker turns out to be the same as in (14), and therefore also  $u_A$  will be the same.

Substituting (17) and (14) into (15), we get this time that:

$$\hat{\alpha}_B = 0.5\mu_B(\theta_0). \quad (18)$$

Hence, since  $\mu_B(\theta_0) < \Pr(\theta_0)$ , the fraction of minority workers promoted is lower than in the corresponding case when their expectations about the

employers' type are correct. Since the fraction of minority workers promoted depends on their beliefs about discriminatory employers and not on the actual fraction, what happens now is that promotions of minority workers do not increase as  $\Pr(\theta_d)$  decreases. Even if discriminatory tastes disappear, i.e.  $\Pr(\theta_d) = 0$ , promotions are still unbalanced across populations as long as minority workers believe they are discriminated against.

This happens because the optimal effort of any worker  $B$  is lower than that of the corresponding type of worker  $A$ . Substituting (18) into (14) we get:

$$e_A^{1*} = 1 + \frac{\Pr(\theta_0)}{\sqrt{\mu_B(\theta_0)}}(\theta_A - 1) > e_B^{1*} = 1 + \sqrt{\mu_B(\theta_0)}(\theta_B - 1).$$

Also in this case all players maximize utility given their beliefs, which are not contradicted by the evidence. In fact, a promotion rate equal to  $\hat{\alpha}_B = 0.5\mu_B(\theta_0)$  is consistent with minority workers' wrong beliefs that the fraction of discriminatory employers is equal to  $\mu_B(\theta_0)$ , thereby defining a fixed point in their learning process. Moreover, plugging (18) into equations (11) and (12), and proceeding with tedious but straightforward computations we obtain that the fraction of minority workers promoted is also in line with what employers and workers  $A$  expect. The uniqueness of this equilibrium refers to the fact that other possible outcomes would violate empiricism, leading to successive updates of beliefs because the learning process would go on.

Hypotheses behind Proposition 1 and Proposition 2 differ only because of expectations of workers  $B$ . In Proposition 1 such expectations are correct, while in Proposition 2 workers  $B$  are assumed to overestimate the percentage of discriminatory employers. Results differ considerably, with wrong expectations of being discriminated against leading to a lower promotion rate among minority workers. Furthermore, nothing changes from the theoretical point of view when it is assumed that  $\Pr(\theta_d) = 0$  (absence of discriminatory employers), meaning that workers' expectations are a "stand alone" source of unequal outcomes from the theoretical point of view.

Minority workers do not "test" their beliefs, meaning that they do not verify whether the employers would have promoted them had they chosen higher effort. The reason is that no single worker has any incentive to experiment, because his observation would have a negligible information value. Only if a sufficiently large fraction of minority workers experiments exerting a higher effort can the initial beliefs be contradicted, but this does not happen because of the classic free-riding problem.

## 4.2 Policy implications

Workers  $B$  are trivially worse off when overestimating the fraction of discriminatory employers, while the opposite holds as far as workers  $A$  are



concerned because of the change in the probability of being promoted that become more favorable to them. Also employers are worse off because workers' productivity is lower on average. In fact, the productivity of minority workers is positively correlated with the fraction of unbiased employers they expect (see Equation 17). Moreover workers  $A$  characterized by high ability do not need to exert an effort higher than the maximum of the best worker  $B$  in order to be sure of being promoted, and such a threshold becomes lower and lower the higher  $\mu_B(\theta_d)$ . Hence, although not Pareto inferior, the equilibrium in Proposition 2 is suboptimal.

Dealing with feedback effects model, Cain (1986) raises a concern which also applies to this model and, more generally, to all the models displaying multiple equilibria some of which suboptimal:

“model’s predicted consequences from a favorable shock are so obviously beneficial to the group discriminated against and to employers that is difficult to see why the upward spiral would not quickly be initiated by group intervention.”

This argument suggests that it should not be difficult to break down unequal outcomes based on workers' expectations, and this is certainly true as far as the mathematics of the model is concerned. However, it is difficult to find a counterpart in the field.

An effective device would be a free insurance that pays back the money equivalent of the utility loss suffered by minority workers who exert a higher effort than the opponent without being promoted, but this has clear problems of observability and therefore enforceability by a third party.

A subsidy to minority workers proportional to their productivity would instead be ineffective in breaking down unbalanced promotions as long as it is common knowledge, and this would be the case for instance of gender based taxation (Alesina, Ichino and Karabarbounis 2009).

Trial work periods are an effective policy tool to break down statistical discrimination outcomes, i.e. a situation in which *employers'* wrong beliefs are self-confirming. In contrast, the equilibria described in Proposition 2 are robust to trial work periods, because they do not affect *workers'* expectations. As long as minority workers think they will be discriminated against, during the trial work period they will still display a lower productivity. In fact, at the end of the first period, which can be regarded as a long trial work period, employers observe workers  $A$  exerting a higher effort on average. Therefore, they promote more workers  $A$  even though there is no bias, either statistical or driven by tastes, against the minority.

Quotas are often implemented to correct unequal outcomes. Despite accomplishing the task of increasing the number of minority workers promoted,

quotas can do it without affecting the causes underneath. The simplest way to implement quotas in this model is to impose that at least a percentage  $q > 0$  of minority workers must be promoted, with  $q$  known by all players. In this model, given that only one worker from each population participates to every constituent game, such a result can be obtained imposing a lottery to the employers. The outcomes of this lottery are that with probability  $q$  employers are forced to promote the minority worker, while with probability  $1 - q$  they are free to choose according to their preferences and updated beliefs. Quotas reduce the importance of effort in determining the outcome, and therefore the effort of all workers will be lower the higher is  $q$ :

$$\begin{aligned} e_A^{1*} &= 1 + (1 - q)Pr(\theta_0)\mu_A(e_B)(\theta_A - 1), \\ e_B^{1*} &= 1 + (1 - q)\mu_B(\theta_0)\mu_B(e_A)(\theta_B - 1). \end{aligned}$$

The fraction of workers  $B$  promoted will increase from  $0.5\mu_B(\theta_0)$  to  $q + 0.5(1 - q)\mu_B(\theta_0)$ , but without affecting minority workers expectations. Once quotas are removed, the fraction of minority workers promoted goes back to the old level.<sup>18</sup> Quotas could be effective in a more general version of the model, when more than one worker from each population participates in the same tournament, provided that quotas induce a competition between minority workers strong enough to make them exert a higher effort. This would cause the rate of promotions within minority workers to increase by more than what directly due to quotas, imposing to the minority workers to update their beliefs toward a lower expected fraction of discriminatory employers.

Summarizing, this model suggests that unequal outcomes can be difficult to tackle when workers' expectations are involved, because many policy tools turn out to be ineffective. The bottom line is that, in line with Coate and Loury (1993), the best way to correct unequal outcomes is to affect expectations of minorities.<sup>19</sup>

## 5 Conclusions

The model presented in this paper analyzes from the theoretical point of view the role of workers expectations, so far neglected in the literature, in explaining the observed unequal outcomes that characterize some minorities in the labour market.

---

<sup>18</sup>A similar trade-off between equity and efficiency associated to affirmative action programs can also be found within uneven tournaments. Experimental evidence, however, do not provide support for such a trade-off, see Schotter and Weigelt (1992) and Corns and Schotter (1999).

<sup>19</sup>For instance, the Gay Pride can be thought as a device that reduces the sexual minorities' expectations of being discriminated against in the labour market.

The model does not explain the rise of discrimination, for which other factors such as discriminatory tastes or statistical discrimination should be held responsible. Even though players have access to the same information, they are likely to interpret the same aggregate observables in different ways, thereby violating the Common Prior assumption. For instance, minority workers can rationalize unequal outcomes via expectations of being discriminated against while majority workers by means of self-serving beliefs concerning the distribution of ability across populations, and employers with statistical discrimination. Although the different explanations for unequal outcomes are very likely to interact among each other, the model has been set up excluding the other causes in order to test whether expectations of being discriminated against stand alone from a theoretical point of view.

The importance of workers expectations can be appreciated comparing the equilibrium outcome in terms of promotions arising when minority workers overestimate the percentage of discriminatory employers with the situation in which such beliefs are correct *ceteris paribus*. Even in a labour market where employers do not discriminate at all against minority workers, and where the distribution of ability is the same across groups of workers, unequal outcomes may still persist because workers expectations of being discriminated against are self-confirming. What happens is that such wrong expectations lower the expected return and therefore the optimal choice of effort of minority workers. Hence, even unbiased employers are more likely to promote majority workers, thereby leading to unequal outcomes that are consistent with initial wrong expectations.

The model illustrates this mechanism within a tournament in which employees compete to be promoted. However, it is possible to interpret, and to implement with minor changes to the theoretical framework, the choice in the first period as the effort workers exert in acquiring signals, or their job search intensity, and employers decision as hiring or assigning applicants to different job types. In this case we would obtain that wrong expectations of being discriminated against are self-confirming in an equilibrium characterized by segregation across jobs of different quality.

Workers' expectation can render many policy actions powerless. For instance, trial work periods, which can be an effective device to break down statistical discrimination outcomes, do not accomplish their task as long as minorities expect to be discriminated against. Affirmative actions like quotas can be successful in increasing the number of minority workers promoted, but this fraction goes back to the old level once they are removed if they do it without affecting workers' expectations.

Hence, expectations of being discriminated against can explain the puzzling long-run persistence of unequal outcomes that characterize some minorities in the labour market, and in particular why historically oppressed social groups are not likely to forge ahead even once the original cause of unequal outcomes has been removed.

## References

- AIGNER, D.J. AND G.G. CAIN (1977): "Statistical Theories of Discrimination in Labor Markets," *Industrial and Labor Relations Review*, 30:2, 175-87.
- ALESINA, A., A. ICHINO AND L. KARABARBOUNIS (2009): "Gender Based Taxation and the Division of Family Chores," *mimeo*.
- ARROW, K.J. (1973): "The Theory of Discrimination" in *Discrimination in Labor Markets* ed. by O. Ashenfelter and A. Rees. Princeton: Princeton University Press, 3-33.
- AUMANN, R.J. (1998): "A Reply to Gul" in *Econometrica*, 66:4, 929-938.
- BATTIGALLI, P. (1987): *Comportamento Razionale ed Equilibrio nei Giochi e nelle Situazioni Sociali*, unpublished dissertation, Universita' Bocconi, Milano.
- BATTIGALLI, P. AND D. GUAITOLI (1997): "Conjectural Equilibria and Rationalizability in a Game with Incomplete Information," in *Decisions, Games and Markets* edited by P. Battigalli, A. Montesano and F. Panunzi. Dordrecht: Kluwer Academic Publishers, 97-124.
- BAYE, M.R., D. KOVENOCK AND C.G. DE VRIERS (1996): "The All-Pay Auction with Complete Information," *Economic Theory*, 8, 291-305.
- BECKER, G.S. (1957): *The Economics of Discrimination*. Chicago: University of Chicago Press.
- (1985): "Human Capital, Effort and the Sexual Division of Labor," *Journal of Labor Economics*, 3, S33-S58.
- BLAU, F.D., M.A. FERBER AND A. E. WINKLER (2006): *The economics of Women, Men and Work*. Upper Saddle River, NJ: Prentice Hall.
- BLAU, F.D. AND C.L. JUSENIUS (1976): "Economists' Approaches to Sex Segregation in the Labor Market: An Appraisal," in *Women and the Workplace* ed. by M. Blaxall and B. Reagan. Chicago: University of Chicago Press, 181-99.
- BREEN R., AND C. GARCIA-PENALOSA (2002): "Bayesian Learning and Gender Segregation," *Journal of Labor Economics*, 20:4, 899-922.
- CAIN, G.G. (1986): "The Economic Analysis of Labor Market Discrimination: a Survey," in *Handbook of Labor Economics*, ed. by O. Ashenfelter and R. Layard. Amsterdam: North Holland, 693-785.

- COATE, S. AND G.C. LOURY (1993): "Will Affirmative-Action Policies Eliminate Negative Stereotypes?," *American Economic Review*, 83:5, 1220-40.
- CORNS, A. AND A. SCHOTTER (1999): "Can Affirmative Action be Cost Effective? An Experimental Examination of Price-Preference Auctions," *American Economic Review*, 89, 291-305.
- DE LA RICA, S., J.J. DOLADO AND C. GARCIA-PEALOSA (2008) "On Gender Wage and Participation Gaps: Theory, Policies and Some Empirical Evidence," *IZA Discussion Paper n.3553*.
- DEKEL, E., D. FUDENBERG AND D.K. LEVINE (2004): "Learning to Play Bayesian Games," *Games and Economic Behavior*, 46, 282-303.
- FERBER, M.A. AND H.M. LOWRY (1976): "The Sex Differential in Earnings: A Reappraisal," *Industrial and Labor Relations Review*, 29, 377-87.
- FILIPPIN, A. (2009a): "A Generalization of the Self-Confirming Equilibrium Concept to Aggregate Observables," *mimeo*.
- (2009b): "Discrimination and Workers' Expectations: a View from the Lab," *mimeo*.
- FILIPPIN, A. AND A. ICHINO (2005): "The Gender Wage Gap in Expectations and Realizations," *Labour Economics*, 12:1, 125-145.
- FUDENBERG, D. AND D.K. LEVINE (1993): "Self-Confirming Equilibrium," *Econometrica*, 61:3, 523-545.
- (1998): *The Theory of Learning in Games*. Cambridge and London: MIT Press.
- GUL, F. (1998): "A Comment on Aumann's Bayesian View" in *Econometrica*, 66:4, 923-927.
- HARSANYI, J.C. (1967-68): "Games with Incomplete Information Played by Bayesian Players," Parts I, II, III *Management Science* 14, 159-182, 320-334, 486-502.
- HOFF, AND P. PANDEY (2004): "Belief Systems and Durable Inequalities," *World Bank Policy Research Working Paper n.3351*.
- LAZEAR, E. AND S. ROSEN (1981): "Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy*, 89, 841-864.
- MINCER, J. AND S. POLACHEK (1974): "Family investments in Human Capital: Earnings of Women," *Journal of Political Economy*, 82:2, 76-108.

- MORO, A. AND P. NORMAN (2002): "A General Equilibrium Model of Statistical Discrimination," *Journal of Economic Theory*, 114:1, 1-30.
- MYRDAL, G. (1944): *An American dilemma*. New York: Harper & Brothers.
- O'KEEFFE, M., W. K. VISCUSI, AND R. J. ZECKHAUSER (1984): "Economic Contests: Comparative Reward Schemes," *Journal of Labor Economics*, 2:1, 27-56.
- PHELPS, E.S. (1972): "The Statistical Theory of Racism and Sexism," *American Economic Review*, 62:4, 659-661.
- SCHOTTER, A. AND K. WEIGELT (1992): "Asymmetric Tournaments, Equal Opportunity Laws, and Affirmative Action: Some Experimental Results," *Quarterly Journal of Economics*, 107, 511-39.
- SINISCALCHI, M. (2007): "Epistemic Game Theory: Beliefs and Types," *The New Palgrave Dictionary of Economics*, forthcoming, Palgrave Macmillan.
- SPENCE, M.A. (1973): "Job Market Signaling," *Quarterly Journal of Economics*, 87:3, 355-374.