Discrimination and Workers' Expectations

Antonio Filippin

July 2003

Discrimination and Workers' Expectations

Antonio Filippin
*University of Milan, EUI
and IZA Bonn*

ABSTRACT

Discrimination and Workers' Expectations[*]

The paper explores the role of workers' expectations as an original explanation for the puzzling long run persistence of observed discrimination against some minorities in the labor market. A game of incomplete information is presented, showing that *ex ante* identical groups of workers may be characterized by unequal outcomes in equilibrium due to their different beliefs, even though discriminatory tastes and statistical discrimination by employers have disappeared. Wrong beliefs of being discriminated against are self-confirming in this circumstance, being the ultimate cause of a lower percentage of promotions which supports these wrong beliefs.

Antonio Filippin
Department of Economics
University of Milan
Via Mercalli, 23
20122 Milan
Italy
Email: antonio.filippin@unimi.it

1 Introduction

Despite several contributions to the literature, there is still no widely shared explanation for the long-run persistence of discrimination in the labor markets. Moreover, the neoclassical theory of discrimination is mostly a demand-side theory. There are very few contributions where workers' heterogeneity matters, and, to the best of my knowledge, only a recent paper by Breen and Garcia-Penalosa (2002) studies the possibility that unequal outcomes may persist for reasons attributable to workers' expectations. The goal of this paper is to set up a static model where preferences and beliefs of both sides of the labor market matter. The advantage of a theoretical framework obtained following this approach is twofold. First, the main contributions to the discrimination literature can be nested and therefore jointly analyzed. Second, it is possible to explore the role of workers' expectations, so far neglected in the literature, which are instead the focus of this paper.

The model is formalized as a two-stage game of incomplete information in which populations of workers and employers are engaged. In every constituent game, i.e. in every repetition of the game played by actors randomly drawn from their populations, three players participate: one employer and two workers, one of whom belongs to a minority group. The employer promotes one (and only one) of the two workers after having observed their output, which is a function of effort and taste for work, the latter unobservable. Crucially for the results of this paper, promotions depend via effort on workers' expectations about the unknown employer's type, which captures the possible disutility of promoting a minority worker. More generally, promotions depend on both employers' and workers' type as well as on their beliefs about opponents' type-strategy profile.

The importance of workers' expectations can be appreciated by comparing the equilibrium outcome in terms of promotions arising when minority workers overestimate the percentage of discriminatory employers as opposed to a situation in which beliefs are correct *ceteris paribus*. Even in a labor market where employers do not discriminate against minority workers either directly or statistically, and where the distribution of ability and taste for work is the same across groups of workers, unequal outcomes may still arise due only to workers' expectations. It is worth stressing that such assumptions are made in order to test workers' expectations as a "stand-alone" source of unequal outcomes from a theoretical point of view, not because other sources are regarded as negligible. What happens is that wrong beliefs of being discriminated against are self-confirming in equilibrium. Minority groups who expect of being discriminated against supply less effort on average, because of a lower expected return. This induces a lower percentage of promotions within minority workers even though employers do not discriminate against them either directly or statistically. In turn, this outcome is consistent with minority workers' beliefs that there are employers characterized by discriminatory tastes.

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, because no single worker has any incentive to experiment. Moreover, the

main result, i.e. that unequal outcomes may be ascribable to workers' different expectations, is robust both to trial work periods, which are instead an effective policy device to break down statistical discrimination outcomes, and to a not too strong degree of affirmative actions like quotas. The conclusion is that workers' expectations may well contribute the puzzling long-run persistency of unequal outcomes observed in the labor market.

The structure of this paper is as follows. After some definitions are outlined in Section 1.1, the constituent game of the model, i.e. the game after the players have already been matched, is presented in Section 2.1. The population game, the matching process and the information structure, necessary to characterize beliefs, are described in Section 2.2. The connections of the model to the related literature are sketched in Section 3. Section 4 concentrates on the analysis of the equilibria of the model and its policy implications. Section 5 concludes.

1.1 Definitions

Before presentating the model, it is useful to clarify the meaning of certain concepts used throughout this paper.

First, **productivity** stands for output per worker. It does not refer to a worker's innate characteristic. In this model productivity is something endogenous, determined by ability, effort and taste for work. Therefore, a more productive worker is simply a worker characterized by a higher output.

Second, in the literature many different and occasionally contradicting definitions have been used referring to discrimination in the labor market. Discrimination has been defined either as different achievements (wages, promotions) for equally *productive* workers, or as different achievements for *ex ante* equal workers, i.e. for workers with the same *ability* and *taste for work*. Not infrequently, the two concepts have been used interchangeably, but this seems inappropriate, because *ex ante* equal workers can be characterized by different productivity in equilibrium.

A good compromise, partially following Blau, Ferber and Winkler (2002), is to use two different definitions. On the one hand, following the "equal pay for equal work" principle, **direct discrimination** can be defined as a different treatment in terms of wages, promotions, or job allocations for equally productive workers.[1] On the other hand, a more comprehensive definition seems to be 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 the fact that many factors, and direct discrimination can be one of the most important, 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 no evidence of direct discrimination. Such a situation is captured by the more comprehensive concept of **cumulative**

---

[1] Often the definition of direct discrimination refers to "equally qualified" workers.

discrimination, defined as different achievements for *ex ante* equal workers.

Another distinction that deserves to be mentioned is that between *group* and *individual* discrimination. The former happens when different achievements are observed on average either between groups of workers that are on average equally productive (direct group discrimination) or between groups of workers which are *ex ante* equal (cumulative group discrimination). The latter happens when an individual is judged on the basis of group membership rather than upon his or her own characteristics only. Individual discrimination is a characteristic of all the models of incomplete information and concerns both the majority and the minority group. Moreover, it does not imply group discrimination. Henceforth, even though not specified, discrimination always refers to group discrimination.

2 The Model

The model is formalized as a two-stage game of incomplete information in which populations of workers and employers are engaged. The two populations of workers differ because of an observable characteristic (race, gender, etc.) which does not affect their output (productivity) π. The observable characteristic distinguishes the so-called majority worker, identified by superscript A, from the so-called minority worker, identified by superscript B. Employers are denoted by superscript 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 2.2.

2.1 The constituent game

In every constituent game one employer and two workers, one of whom is a "minority" worker, are drawn from their populations and play a two-stage game. In the first period both workers choose one out of three possible levels of effort $e_A^1, e_B^1 \in E = \{l, i, h\}$, where l stands for "low", i stands for "intermediate" and h stands for "high". The employer observes workers' productivity in the first period and promotes one (and only one) of the two workers. After having observed the employer's decision, workers choose a level of effort for the second period $e_A^2, e_B^2 \in E$.

The constituent game is characterized by observable actions, because the decision about promotion is directly observed and every level of (observed) output is one-to-one related to effort.[2] The game is also characterized by incomplete information, because every player knows his or her type only.

2.1.1 Incomplete information

In this game every player knows his\her own type only.

---

[2] More precisely: $\pi = e$. See assumption 1 below.

$\theta_A \in \Theta_A$ and $\theta_B \in \Theta_B$ summarize the type of majority and minority workers, respectively. Workers' type represents their taste for work, formalized as a weight attached to the disutility of effort in the second period (see equations (1) and (2) below). There are only two possible types of worker $\Theta_A = \Theta_B = \{1, K\}$. The lower the parameter, the lower the cost of effort.[3]

$\theta_F \in \Theta_F$ represents employer's tastes for discrimination. There are only two types of employer $\Theta_F = \{0, d\}$, with d sufficiently large. If $\theta_F = 0$ the employer is indifferent about the observable characteristic which distinguishes the workers. On the other hand, if $\theta_F = d$ the employer suffers a disutility when the minority worker is promoted. The disutility d is assumed to be sufficiently high that promoting worker A is always the optimal choice regardless of workers' productivity.

Summarizing, a minority worker knows her own taste for work θ_B, while the type θ_A of the majority worker and the tastes for discrimination θ_F of the employer are unknown.[4]

2.1.2 Payoffs

The structure of the utility function is the same for majority (A) and minority (B) workers and it is parametrized according to their type θ. The analysis focuses on risk neutral workers, whose lifetime utility is

$$U^{\theta_A} = w_A^1 - (e_A^1)^2 + w_2^A - \left(\alpha \frac{\theta_A}{K} + (1 - \alpha)\theta_A \right)(e_2^A)^2, \qquad (1)$$

$$U^{\theta_B} = w_B^1 - \left(e_B^1 \right)^2 + w_2^B - \left((1 - \alpha)\frac{\theta_B}{K} + -\alpha\theta_B \right)\left(e_2^B \right)^2, \qquad (2)$$

where:

w_P^t is the wage in period t for the worker belonging to population $P = (A, B)$.

e_P^t is effort in period t for the worker belonging to population $P = (A, B)$.

$\alpha = 1$ if worker A is promoted; $\alpha = 0$ if worker B is promoted.

$K > 1$ summarizes that the job assigned to the promoted worker is more desirable, because for any given level of effort the disutility will be lower.

Assumption 1. $w^t = \pi^t = e^t$. The labor market is assumed to be competitive, therefore productivity is entirely paid to workers. Moreover, productivity coincides with effort. This assumption makes the game equivalent to the reduced form of a more general game where workers' output is observed and verifiable

---

[3] Taste for work is assumed to matter in the second period only, as if one of the two types of worker ($\theta = K$) became lazier after the promotion decision. Allowing different tastes for work in the first period would *de facto* resolve an employer's uncertainty about workers' type before the decision about promotion is taken. Such an uncertainty is instead necessary to get the results that are shown in section 4. To get the same results even relaxing this assumption it would be necessary to remove the restriction that ability is equal for all workers. A more plausible model would be obtained providing the same insights at the price of a substantially increased complication (see also footnote 5).

[4] Of course, the same holds *mutatis mutandis* for players B and F.

and employers compete on enforceable piece-rate contracts. Workers would be free to move, but in equilibrium $w^t = e^t$ and nobody moves.[5]

It follows from the utility function in (1) and (2) that in the second period effort will be higher if the worker is promoted

$$
\begin{aligned}
e_A^{2*}|\,(\alpha = 1) &= \frac{K}{2\theta_A} > e_A^{2*}|\,(\alpha = 0) = \frac{1}{2\theta_A}, \\
e_B^{2*}|\,(\alpha = 0) &= \frac{K}{2\theta_B} > e_B^{2*}|\,(\alpha = 1) = \frac{1}{2\theta_B}.
\end{aligned}
$$

Substituting the type of workers $\theta_A, \in \theta_B\{1, K\}$ into these equations, it becomes evident that a bad type who is promoted supplies the same effort as a good type who is not promoted.[6]

Assumption 2: *The set of levels of effort is* $E = \left\{l = \frac{1}{2K}, i = \frac{1}{2}, h = \frac{K}{2}\right\}$. Both in the first and in the second period there are only three conceivable levels of effort $h > i > l > 0$ that stand for "high," "intermediate" and "low" and that coincide with the optimal choice in the second period of a bad type who is not promoted, of a bad type who is promoted as well as of a good type who is not promoted, and of a good type who is promoted, respectively.

As far as the employer is concerned, the utility function contains both profits and a parameter summarizing the disutility associated with the promotion of workers B. This means that only workers B face the risk of being discriminated against, because of the observable characteristic that, without affecting their productivity, differentiates them from workers A. Since productivity is assumed to be entirely paid to workers, in this model discrimination can only assume the form of denying a promotion to a worker B that would deserve it. Employer's utility function is

$$
U^{\theta_F} = (m-1)(\pi_A^1 + \pi_B^1 + \pi_A^2 + \pi_B^2) - (1-\alpha)\theta_F,
$$

where $m > 1$ is a known and constant mark-up on workers' productivity due to the entrepreneurial activity. Therefore, in order to maximize profits, the employer needs to maximize workers' productivity. In other words, the employer has an incentive to promote the more productive worker.[7] The term $(1-\alpha)\theta_F$ represents the disutility associated with the promotion of a minority worker. When $\theta_F = 0$ the observable characteristic that distinguishes the workers does not matter and profits are the only thing that the employer considers. On the other hand, when $\theta_F = d$ the employer is characterized by discriminatory tastes.

---

[5] The assumption that $\pi = e$ implies that ability does not matter in this model. A more general version in which ability varies across workers but is identically distributed within populations turns out to be much more complicated without being more insightful (see also footnote 3).

[6] The specification of the utility function adopted in the paper is the same as that proposed by the asymmetric tournament literature (see O'Keeffe, Viscusi, and Zeckhauser (1984)). The only difference is that here the role of the prize is played by the lower cost of effort attached to the job assigned to the promoted worker.

[7] It is also possible to interpret F as a supervisor, which maximizes a bonus that is a fraction of the overall productivity of the workers. Nothing would change, because also the supervisor has the incentive to promote the more productive worker in order to maximize his\her bonus.

2.1.3 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. For both populations of workers effort can take the same three values only: $e^1, e^2 \in \{h, i, l\}$.

The employer observes each worker's productivity in the first period and then promotes one (and only one) of them to a more rewarding position. The set of feasible actions for the employer, regardless of his\her type, is simply $\alpha = \{0, 1\}$, where $\alpha = 1$ stands for "promotes worker A" and $\alpha = 0$ stands for "promotes worker B." As far as the employer is concerned, strategies s_F are therefore a vector that specifies a promotion decision for every possible pair of observed productivity levels.

Assumption 3: *Strategies of non-discriminatory employers implement a fair contest.* In other words, strategies of non-discriminatory employers are invariant to the permutation of the observable characteristics that distinguishes workers, *ceteris paribus*, meaning that promotions depend on productivity only (see Lazear and Rosen, 1981). The rationale is to prevent unequal outcomes from arising because of asymmetric choices of employers who are instead supposed to be indifferent to workers' membership.

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

2.2 The Population Game

The constituent game described in section 2.1 is inserted in a wider game, called the population game, which specifies how players are matched and what information they can access. The description of the information structure makes it possible to define players' beliefs.

There are three populations, one of employers and two of workers. As already said for the constituent game, the two populations of workers differ because of an observable characteristic (gender, race, etc.) that does not affect their productivity. The distribution of types within the two populations of workers is identical. This assumption rules out the possibility that unbalanced promotions across populations arise because of a different average disutility of work.

2.2.1 Matching

Each of the three populations $P = \{A, B, F\}$ is composed of a continuum of players, so that the law of large numbers can be invoked to ensure that the actual fraction of any given combination of types of A, B and F players coincides (almost surely) with its probability. The three populations play an infinitely repeated game. At every stage each player of population A is randomly matched with one player from population B and one player from population F.

2.2.2 Information structure

Players try to figure out the distribution of types and strategies among the populations of opponents using the available information. Besides observing the terminal history, i.e. the complete sequence of actions, $z = (w_A^1, w_B^1, \alpha, w_2^A, w_2^B) \in Z$ of the constituent game in which she is involved, every player also observes the distribution of terminal histories $\sigma \in \Delta(Z)$, i.e. the distribution of complete sequences of actions that has occurred in all the constituent games. Players are therefore assumed to access very informative statistics: individual outcomes are a useless source of information when compared to available aggregate outcomes. The availability of aggregate outcomes under observable actions rules out the possibility that unequal outcomes arise in equilibrium, like in Breen and Garcia-Penalosa (2002), as an inheritance of past differences in fundamentals.[8]

2.2.3 Beliefs

Beliefs of a player are a probability measure over the unknown component of the type-strategy set $\Theta \times S = \Theta_A \times \Theta_B \times \Theta_F \times S_A \times S_B \times S_F$. Given that every player is supposed to know his\her own type and the strategy he\she chooses only, the unknown component of $\Theta \times S$ turns out to be the set of type-strategy profiles of all the other players, both the opponents and the other players of his\her own population. Beliefs of a worker of population A when his\her type is θ are defined

$$\mu^{\theta_A} \in \; \Delta\left(\Theta \times S\right).$$

Assumption 4: *Beliefs exclude the possibility that opponents correlate their play:*

$$\mu^{\theta_A}(\theta_A, s_A, \theta_B, s_B, \theta_F, s_F) = \mu^{\theta_A}(\theta_A, s_A)\mu^{\theta_A}(\theta_B, s_B)\mu^{\theta_A}(\theta_F, s_F).$$

Since every player knows his\her own type and strategy, the appropriate marginal distribution for worker A is

$$\mu^{\theta_A}(\theta_B, s_B, \theta_F, s_F) = \mu^{\theta_A}(\theta_B, s_B)\mu^{\theta_A}(\theta_F, s_F).$$

Something more needs to be said about employers' beliefs. Before deciding, the employers have the opportunity to update their beliefs observing workers' productivity. Employers' prior beliefs are a probability measure over each

---

[8] Under a more general setting of the model, e.g. when ability also plays a role, observing the joint distribution of histories would be too informative for the Self-Confirming Equilbria of Section 4 to survive. In that more general setting, only the marginal distribution across populations should be observed for the Self-Confirming Equilibria to survive, but then the individual information would be more informative, allowing players to compute the probability to be promoted conditional on workers'output, while aggregate information does not. For the aggregate information to be still more informative, it would be necessary that each worker participate in the labor market for a finite and not too big number of rounds.

worker's type-strategy profile $\mu^{\theta_F}\left(\theta_A, s_A\right), \mu^{\theta_F}\left(\theta_B, s_B\right)$. Such beliefs can be revised independently using Bayes rule, given that the productivity of worker A does not convey information about worker B and vice versa. Defining

$$\mu^{\theta_F}(\pi_A^1) = \sum_{(\theta_A, s_A) \in (\Theta_A \times S_A)} \mu^{\theta_F}\left\{\theta_A, s_A : e_A^1 = \pi_A^1\right\}$$

the probability that an employer of type θ assigns to the productivity level $\pi_A^1 \in \Pi_A$ according to his\her prior beliefs, updated beliefs after the observation of a productivity level $\tilde{\pi}_A^1$ will be:

$$\mu^{\theta_F}(\theta_A, s_A | \tilde{\pi}_A^1) \mu^{\theta_F}(\tilde{\pi}_A^1) = \left\{ \begin{array}{ll} \mu^{\theta_F}(\theta_A, s_A) & \text{if } \theta_A, s_A : \pi_A^1 = \tilde{\pi}_A^1 \\ 0 & \text{if } \theta_A, s_A : \pi_A^1 \neq \tilde{\pi}_A^1 \end{array} \right. .$$

Although this model does not deal with dynamics, I think it is useful to provide an intuition about how beliefs may be formed in this game. Beliefs of a player at time t can be thought to be a function of the available information about aggregate outcomes arising from the previous period σ_{t-1}. Notice that the same sequence of observables can lead to different beliefs. In other words, players can interpret in different ways the same information about aggregate outcomes. For example, workers can interpret a given distribution of promotions across populations A and B assigning different weights to the role played by workers' effort and employers' propensity to discriminate against the minority. Of course, asymptotic empiricism requires that in equilibrium all the belief rules must generate subjective distributions of observables which coincide with the objective one. For instance, the dataset used by Filippin and Ichino (2003) provides interesting evidence that men and women share very similar expectations about the magnitude of the gender wage gap, arguably because they access the same information about the realized gender gap. However, they assign a different importance to the underlying causes. In fact, while a larger fraction of men think that "actual differences between men and women" matter, a larger fraction of women points towards the "employers' discriminatory tastes" as one of the causes for the expected gap.

Assumption 5: *Workers believe that employers' strategies are weakly monotone in productivity:*

$$\begin{aligned} \forall e_B^1, \quad \mu\left(\alpha = 1 | h, e_B^1\right) &\geq \mu\left(\alpha = 1 | i, e_B^1\right) \geq \mu\left(\alpha = 1 | l, e_B^1\right) \\ \forall e_A^1, \quad \mu\left(\alpha = 0 | e_A^1, h\right) &\geq \mu\left(\alpha = 1 | e_A^1, i\right) \geq \mu\left(\alpha = 1 | e_A^1, l\right) \end{aligned}$$

Workers think that the probability of being promoted cannot decrease when effort increases *ceteris paribus.* This assumption is a way of refining the set of equilibria. The intuitive reason is that promotions are desirable, and workers may be willing to give up utility in the first period in order to enhance their probability of being promoted. The way to do this is to deviate from effort i in the first period, where i is the optimal choice in a world without promotions, supplying either l or h. Assumption 5 means that one's willingness to be promoted will be signaled only through a higher effort. All the equilibria that could

possibly arise when workers signal their willingness to be promoted supplying l are excluded. Moreover, exerting effort l in the first period turns out to be strictly dominated for all workers.

Before characterizing the equilibria of the game presented so far, it is useful to review briefly some of the contributions to the discrimination literature.

3   Related Literature

The model presented so far is flexible enough to capture, under appropriate assumptions, the main features of most of the contributions to the discrimination literature.[9] One thing that must be taken into account is that, although these contributions often focus on wages rather than promotions, the main stylized facts can be replicated focussing on promotions as well.

Six groups of models are presented: discriminatory tastes, statistical discrimination, human capital theory, feedback effects, workers' expectations and asymmetric tournaments.

3.1   Discriminatory tastes

The starting point of the economic analysis of discrimination in labor markets can be found in 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, co-workers 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, co-workers or salespeople.

The following are the assumptions that must be imposed on the model presented in Section 2.1 and 2.2, in order to make it equivalent to the discriminatory tastes approach.

1. The employer's type set is a singleton $\Theta^F = \{d\}$, with $d > 0$.

2. Beliefs assign probability one to the true type-strategy profile of the opponents (absence of uncertainty). In other words, expectations do not matter.

---

[9] In this section, theories have been selected and outlined in such a way as to facilitate contrast and comparison with the model of workers' expectations. Therefore, the choice of the contributions is far from being exhaustive, focussing only 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 about what has been considered more important in the literature thus far. Another reason for these choices is that many detailed surveys are already available (see Blau, Ferber and Winkler (2002) and Cain (1986) among others).

While in Becker's model discrimination takes the form of different pay for equal work, in the game obtained imposing these assumptions discrimination takes the form of always promoting worker A.

Among the advantages of Becker's approach, there is the possibility of explaining the rise of any type of direct discrimination (based on sex, race, religion, etc.). On the other hand, 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. The reason is that discrimination is costly for the employer, so that when competition drives profits toward zero discriminatory employers will suffer a negative utility. Alternatively, we should observe complete segregation. However, both predictions are contradicted by empirical evidence.

3.2 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. Common to these models is the fact that, unlike Becker's one, fundamentals are not relevant.

The most representative model of statistical discrimination has been proposed by Arrow (1973).[10] Employer's beliefs about the existence of different characteristics between (*ex ante* identical) groups turn out to be correct in equilibrium.[11] Why are these expectations confirmed in equilibrium even if the groups were equal *ex ante*? In other words, why are these 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 investment in his/her skills (or a suboptimal supply of effort) and therefore determining an outcome that confirms the beliefs of the employer. The conclusion is that in equilibrium there is *cumulative but not direct discrimination*, because workers are *ex ante* equal but show a different productivity in equilibrium.

Statistical discrimination outcomes, as modeled by Arrow, can be replicated within the game of section 2.1 and 2.2, provided that appropriate changes are made to the structure of the constituent game. Such changes are necessary because in Arrow's model the employer plays using prior beliefs, contrarily to the game presented above. Promotions have to be modified into job assignment

---

[10] Other examples of statistical discrimination 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 pioneering work about signaling. A skeptical reading of the statistical discrimination approach can be found in Aigner and Cain (1977) and Cain (1986). Some of the arguments raised by Cain are also relevant in the context of the model of workers' expectation presented in this paper and are therefore explicitly addressed in what follows.

[11] Moro and Norman (2002) analyze statistical discrimination using a general equilibrium approach.

decisions for the two models to be equivalent. Hence, the whole first period has to be cancelled and appropriate assumptions have to be made accordingly.

1. Employers' strategies coincide now with feasible actions $\alpha = \{0, 1\}$, with $\alpha = 1$ standing for "assign worker A to the good job and worker B to the bad job" and vice versa for $\alpha = 0$.[12] Workers' strategies contain two effort levels, one if assigned to the good job and another if assigned to the bad job.

2. Payoffs become:

$$U^{\theta A} = w_A - (1 - \alpha + \frac{\alpha}{K})\theta_A \left(e_A\right)^2,$$

$$U^{\theta B} = w_B - (\frac{1 - \alpha}{K} + \alpha)\theta_B \left(e_B\right)^2,$$

$$U^{\theta F} = (m - 1)(\pi_A + \pi_B) - (1 - \alpha)\theta_F.$$

3. Discriminatory tastes do not play any role, i.e. the set of employer types is a singleton $\Theta^F = \{0\}$. The employer's action is also observed,. The implication of this is that workers' expectations do not matter any more, because when they move their uncertainty has already been resolved.[13]

4. Employers believe that minority workers are on average less productive in the good job. Defining as $\mu^{\theta_F}(\pi_A | \alpha = 1)$ the employer's expectation about productivity of worker A if assigned to the good job, statistical discrimination means that the average productivity if assigned to the good job is thought to be lower for workers of population B

$$\sum_{\pi_A \in \Pi_A} \mu^{\theta_F}(\pi_A | \alpha = 1) \pi_A > \sum_{\pi_B \in \Pi_B} \mu^{\theta_F}(\pi_B | \alpha = 0) \pi_B.$$

Employers expect minority workers to be less productive, and therefore assign them to the bad job. Worse job assignment causes minority workers to exert a lower effort, with the result that *ex post* they are actually less productive, confirming employer's expectations.

Statistical discrimination models have been criticized by Cain (1986), on the ground that "these models face the criticism that the employer's uncertainty about the productivity of workers may be inexpensively reduced by observing the workers' on-the-job performance." Workers' performance can be observed, for example, by means of trial work periods. Cain's argument can straightforwardly be encompassed into the model presented in this paper going back to the original version of the constituent game, where updated beliefs are used to decide on promotions and where the whole first period can be thought as a trial

---

[12] The good and the bad job coincide with the job assigned in section 2.1 to the promoted and to the non-promoted worker, respectively.

[13] Fryer (2002) presents a dynamic model of statistical discrimination where workers' beliefs matter but are assumed to be correct.

work period. Nonetheless, the statistical discrimination model plus trial work period leaves some open questions: what determines workers' behavior in the trial work period? Is it convenient for them to increase effort to be assigned to the good job? The answers to these questions cannot be found within the statistical discrimination literature, because it is necessary to analyze also the supply side of the labor market. In section 4, where the role of workers' expectations is analyzed, 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.

3.3   The human capital theory

Another strand of the literature, started by Mincer and Polacheck (1974), is the so-called human capital theory which 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 labor force. As a consequence, they receive less on-the-job training and/or are assigned to less rewarding jobs. Such behavior can be ascribed to the traditional division of work within the family (Becker, 1985). In this way, wage differentials, worse career path, and/or sex segregation are explained by voluntary choices. If this is the case, the different achievements *could not be classified as discrimination*, given that workers are neither equally productive in equilibrium nor *ex ante* equal.

The human capital theory can easily be nested into this model assuming that:

1. $\Pr(\theta_A = 1) > \Pr(\theta_B = 1)$. Having a higher average disutility of effort, minority workers find it optimal to supply on average a lower effort. Hence, they are less productive.

A criticism that some economists have made of this approach (see the next subsection) is that the seemingly "voluntary" decision could actually be induced by discrimination, entering the definition of cumulative discrimination.

3.4   Feedback effects

The boundaries of this approach are particularly uncertain,[14] and usually surveys concerning discrimination in the labor market use these models as a coun-

---

[14] A large number of the so called "institutional" contributions may also fall into this category. The seminal "institutional" contribution has been made by Myrdal (1944), who theorizes the "principle of cumulation," a mechanism of dynamic causation between several variables. These variables move together influencing each other once the system is hit by an external shock. Among the secondary causes of discrimination, the behavior of workers is also taken explicitly into account: "The Negro worker often feels that his fate depends less on his individual efforts than on what white people believe about Negroes in general" (Myrdal, 1944). Other contributions follow along the line of the vicious circle described by Myrdal, like Ferber and Lowry (1976).

terpart for other theories, without analyzing them separately. The reason is that the contributions that can be grouped into this category are quite heterogeneous: the main idea they have in common is that the behavior of the workers can in turn be determined by discrimination. However, the mechanisms through which the behavior is affected vary considerably. In many cases there is also a lack of formalization and these effects are little more than qualitative statements.

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.

No specific assumption is necessary to nest this approach into the game of section 2.1 and 2.2, because the presence of workers' expectation is *per se* a way of formalizing such feedback effects.

3.5 Workers' expectations

As already mentioned, the neoclassical theory of discrimination is mostly a demand-side theory. But why should workers' expectations not be allowed to play a role as important as that of either employers' preference in the discriminatory tastes approach or employer's beliefs in the statistical discrimination models?

To the best of my knowledge, the only paper in the literature on discrimination that focuses on the supply side of the labor 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 cause 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.

Similarities of this paper with the work of Breen and Garcia-Penalosa are evident: both consider the effect of heterogeneity within the supply side of the labor market and both explain the persistence of unequal outcomes using self-confirming workers' expectations. Besides a dynamic setting that prevents it from being nested within the framework of Section 2.1 and 2.2, what differs in the model of Breen and Garcia-Penalosa is a different information structure. Agents learn from their parents only, and 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 key mechanism behind the results of these authors, is that the information structure of

the model prevents agents from learning that differences in fundamentals have disappeared. In other words, beliefs are still a function of differences in workers' fundamentals. Section 4 will show that when such an assumption is relaxed within a static framework, workers' expectations can still explain observed unequal outcomes.

3.6 Asymmetric tournaments

The literature on tournaments, started by Lazear and Rosen (1981), is not directly related to discrimination. Nevertheless, asymmetric tournaments, as described by O'Keeffe, Viscusi, and Zeckhauser (1984), provide a useful framework for the analysis of the effects of discrimination on promotions. Therefore, asymmetric tournaments are a natural and valuable benchmark for the game presented in this paper. As already mentioned, 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 favor one of them.

Within the literature on uneven tournaments, it is incidentally mentioned that unequal outcomes may arise between *ex ante* equal groups of workers.[15] However, the underlying mechanism has not been formalized and, more specifically, no role is explicitly played by expectations. Another difference with respect to this model is that effort is continuous and imperfectly observable. Moreover, there is only one period and the winner is determined by the largest drawing of effort together with possible handicaps imposed by unfair rules. Predictions that emerge throughout this paper will be compared with the corresponding predictions coming from asymmetric tournaments. Not surprisingly, a lot of similarities arise.

4 Analysis of the equilibria

In this section two qualitatively different sets of equilibria are presented. The first set (see Proposition 1) displays symmetric outcomes under the assumption that the expectations of all players are correct. The second set of equilibria (see Proposition 2) shows that unequal outcomes may arise when minority workers' expectations are wrong *ceteris paribus*.

Two different concepts are necessary to analyze the equilibria of the model: the Self-Confirming Equilibrium (henceforth: SCE) as generalized in Filippin (2003a) to cover the case of aggregate observables, and the Bayes-Nash Equilibrium (henceforth: BNE). The two concepts share the feature that each player maximizes utility given his\her beliefs, updated whenever possible, about every possible opponents' type-strategy profile (see section 4.1). The difference between them is that in a BNE each player has a correct conjecture about the relationship between opponents' types and choices, i.e. about their behavioral rules. In the commonly applied subcase when beliefs satisfy the Common Prior

---

[15] See Schotter and Weigelt (1992).

assumption, beliefs about opponents' types are correct as well. On the other hand, when the Common Prior assumption is not satisfied, beliefs in a BNE may be contradicted by the evidence. On the contrary, in an SCE beliefs may not coincide with the true distribution of opponents' type-strategies profile, as long as they are not contradicted by the evidence (see Battigalli and Guaitoli (1997) and Dekel, Fudenberg and Levine (2003) for a formal discussion of the relation between the Common Prior assumption, BNE and SCE in games of incomplete information).

The equilibria in both Proposition 1 and Proposition 2 are Perfect Bayes-Nash (henceforth: PBNE) and Self-Confirming at the same time. This is fairly intuitive in Proposition 1 given that the Common Prior assumption is satisfied and therefore beliefs are correct. Nevertheless, this is the case also in Proposition 2 even though the Common Prior is violated. In fact, neither beliefs of worker A nor beliefs of worker B, although different, are contradicted by the evidence. Furthermore, beliefs of both workers correctly predict players' behavioral rules (see section 4.2).

4.1 Utility maximization given beliefs

a) Employers

Only workers' difference in productivity *after* the promotion affects the employer's decision, while the difference in the first period does not. The reason is that the disutility θ_F is associated with the *promotion* of a minority worker. Therefore, at the margin only benefits from the promotion of a minority worker (i.e. difference in productivity after promotion) are compared with the associated cost θ_F in order to decide which worker is optimal to promote.

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. Hence, they will always promote the worker they think will be more productive *after* the promotion, regardless of the population she comes from. If workers are of the same type the overall productivity and the employer's utility after the promotion are the same regardless of the worker who is promoted. On the other hand, if workers' type is different the employer maximizes his\her utility promoting the worker characterized by higher taste for work (i.e. lower θ).

Defining $\mu^{0_F}(\pi_A^2|\pi_A^1)$ as the updated beliefs of a non-discriminatory employer about the productivity π^2 of worker A in the second period having observed π_A^1 in the first, it follows that the best reply $BR^{0_F}(\pi^1)$ to the observed pair of productivity levels $\pi^1 = (\pi_A^1, \pi_B^1)$ will depend on the comparison of workers' expected productivity in the second period. Formally:

$$\alpha = BR^{0_F}(\pi^1) = 1 \text{ if}$$
$$\sum_{\pi_A^2 \in \Pi_A} \mu^{0_F}(\pi_A^2|\pi_A^1)\pi_A^2 > \sum_{\pi_B^2 \in \Pi_B} \mu^{0_F}(\pi_B^2|\pi_B^1)\pi_B^2, \tag{3}$$

15

which means that promoting a majority worker is the best reply whenever the majority worker is expected to be strictly more productive in the second period, given the observed productivity levels. Similarly, promoting a minority worker is the best reply whenever equation 3 holds with reversed inequality sign. If expected productivity in the second period is the same, the non-discriminatory employer is indifferent. This means that both $\alpha = 1$ and $\alpha = 0$ as well as all the mixed strategies would be best replies. However, the only non-trivial strategy that implements a fair tournament (assumption 3) and that satisfies monotonicity (assumption 5) is:

$$i, i \to 0.5; \quad i, h \to 0; \quad h, i \to 1; \quad h, h \to 0.5, \tag{4}$$

where the action is defined as "percentage of workers A promoted" and, for example, "h, i" means that the productivity levels of worker A and worker B are *high* and *intermediate*, respectively.[16],[17]

Employers of type $\theta_F = d$ are characterized by tastes for discrimination and, since by assumption d is sufficiently large, they promote worker A regardless of any observed and expected productivity level of the two workers:

$$\alpha = BR^{d_F}(\pi^1) = 1.$$

b) Workers

Workers' optimal actions in the second period according to their type and the employer's decision have already been derived when describing assumption 2. Substituting such values in the utility functions (1) and (2) we obtain:

$$
\begin{aligned}
U &= e^1 - (e^1)^2 + \frac{\alpha K}{4} + \frac{1-\alpha}{4} && \text{for } \theta_A = 1, \\
U &= e^1 - (e^1)^2 + \frac{\alpha}{4} + \frac{K}{4}(1-\alpha) && \text{for } \theta_B = 1, \\
U &= e^1 - (e^1)^2 + \frac{\alpha}{4} + \frac{1-\alpha}{4K} && \text{for } \theta_A = K, \\
U &= e^1 - (e^1)^2 + \frac{\alpha}{4K} + \frac{1-\alpha}{4} && \text{for } \theta_B = K.
\end{aligned}
$$

As far as the first period is concerned, l can easily be shown to be a strictly dominated action for all workers as long as workers' beliefs about employers' strategies are monotone (assumption 5). The utility of a type $\theta_A = 1$ choosing i and h in the first period is, respectively:

$$
\begin{aligned}
U(i) &= \frac{1}{4} + \mu\left(\alpha = 1 | i\right)\frac{K}{4} + \left(1 - \mu\left(\alpha = 1|i\right)\right)\frac{1}{4}, \\
U(h) &= \frac{K}{2} - \frac{K^2}{4} + \mu\left(\alpha = 1 | h\right)\frac{K}{4} + \left(1 - \mu\left(\alpha = 1|h\right)\right)\frac{1}{4},
\end{aligned}
$$

---

[16] Pairs of productivity levels where effort l is involved are not considered because in the first period l is strictly dominated for both workers under assumption 5.

[17] When non-discriminatory employers only are involved, the game becomes equivalent to an all-pay auction, insofar as the utility loss suffered by a non-promoted worker who chose effort h instead of i is sunk; see Baye, Kovenock and de Vries (1996).

where $\mu\left(\alpha=1|i\right)$ and $\mu\left(\alpha=1|h\right)$ are the subjective probabilities to be promoted when playing i and h, respectively.[18] Therefore, type $\theta_A = 1$ will choose h in the first period if $U(h) - U(i) > 0$, which leads to

$$\mu\left(\alpha=1|h\right) - \mu\left(\alpha=1|i\right) > (K-1).$$

Similarly,

- a worker $\theta_B = 1$ will choose h if $\mu\left(\alpha=0|h\right) - \mu\left(\alpha=0|i\right) > (K-1)$;

- a worker $\theta_A = K$ will choose h if $\frac{1}{K}\left(\mu\left(\alpha=1|h\right) - \mu\left(\alpha=1|i\right)\right) > (K-1)$;

- a worker $\theta_B = K$ will choose h if $\frac{1}{K}\left(\mu\left(\alpha=0|h\right) - \mu\left(\alpha=0|i\right)\right) > (K-1)$.

Note that when there is no chance of being promoted, the left-hand side of these equations vanishes and h can never be an optimal choice since $K > 1$.

4.2 Correctness of beliefs

Beliefs are *correct* whenever, for all the players of every type θ of each population, the subjective probability distribution over opponents' type-strategy set coincides with the objective distribution. For instance,

$$\mu^{\theta_A}\left(\theta_B, s_B, \theta_F, s_F\right) = \Pr(\theta_B, s_B, \theta_F, s_F) \tag{5}$$

intuitively means that the probability assigned by players of type θ of population A to every combination of opponents' type-strategy profile is correct.

Beliefs are *not contradicted by the evidence* whenever all the players' subjective probability to observe a particular distribution of histories σ coincide with its actual frequency. The subjective probability to observe σ is obtained summing up the probabilities attached to every combination of opponents' type-strategy profiles that would lead to a combination of observables equal to σ. It may happen that incorrect beliefs, i.e. beliefs which violate (**??**) for some type θ or strategy s, are not contradicted by the evidence.[19]

4.3 Existence of the equilibria

This section focuses on the role of workers' expectations. Appropriate assumptions are imposed in order to neutralize all the other potential causes of unequal outcomes. In particular:

Assumption 6: *Only minority workers' beliefs about employers' type may be wrong, while all the other beliefs are correct.* In particular,

a) $\mu^F(\cdot) = \mu^A(\cdot) = \Pr(\cdot)$ beliefs of both employers and workers A are correct;

---

[18] In $\mu(\alpha|\cdot)$, the superscript identifiyng the type and population of the player is omitted in order to avoid a heavy notation.

[19] The intermediate case, when beliefs are correct only as far as the behavioral rules are concerned, can be represented in the following way:
$\mu^{\theta_A}\left(s_B, s_F|\theta_B, \theta_F\right) = \Pr(s_B, s_F|\theta_B, \theta_F)$ and $\mu^{\theta_A}(\theta_B, \theta_F) \neq \Pr(\theta_B, \theta_F)$.

b) $\mu^B(\theta_A, s_A) = \Pr(\theta_A, s_A)$; $\mu^B(\theta_B, s_B) = \Pr(\theta_B, s_B)$ beliefs of workers B concerning the type-strategy profile of both workers A and the other workers B are correct;

c) $\mu^B(s_F|\theta_F) = \Pr(s_F|\theta_F)$ beliefs of workers B about employers' behavioral rules are correct.

Assumption 7: *All workers B share the same beliefs about employers' type.* This assumption is crucial for unequal outcomes to be produced by workers' expectations within this extremely simplified version of the model. In fact, a small fraction of minority workers with correct beliefs would be enough to falsify the wrong beliefs of the other workers of that population. This would be a serious problem if the goal of the model was to claim that workers' wrong expectations of being discriminated against are *the only* explanation for observed unequal outcomes. On the contrary, less ambitiously as well as much more realistically, the model is simply aimed at stressing that workers' expectations can play a relevant role. Such a goal is pursued in a way that isolates the role of workers' expectations as much as possible. Not surprisingly, equilibria are not robust to some perturbations like that implied by the relaxation of assumption 7, unless some additional degrees of freedom are allowed. This can be done by, for instance, allowing beliefs of workers B about some type-strategy profiles to differ as well, i.e. relaxing assumption 6b at the same time, or allowing more heterogeneity of fundamentals within the model.

Similarly, some of the other assumptions made so far are very convenient from the theoretical point of view, because they allow us to neutralize other causes of unequal outcomes. For instance, the assumption that the distribution of types is the same across populations of workers excludes any role of the human capital approach, while assumption 6a rules out statistical discrimination outcomes.[20] It deserves to be stressed once more that such assumptions are made with the only purpose to focus the theoretical analysis on the role of workers' expectations and not because the other causes of unequal outcomes are regarded as less important.

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 $\mu^B(\theta_F) = \Pr(\theta_F)$ or wrong $\mu^B(\theta_F) \neq \Pr(\theta_F)$, where $\Pr(\theta_F)$ is the distribution of types within the population of 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, a Perfect Bayes-Nash Equilibrium always exists where:*

1) in the first period both types of population A choose the same actions of the corresponding type of population B.

2) the percentage of workers A promoted is equal to $1 - 0.5\Pr(\theta_F = 0)$.

---

[20] The effect of discriminatory tastes has not been neutralized because it is straightforward to see what happens with or without imposing $\Pr(\theta_F = 0) = 1$ in Propostions 1 and 2.

When, in addition to assumption 6, also expectations of workers B about employers' type are correct, $\mu(\alpha|\cdot)$ can be substituted with $\Pr(\alpha|\cdot)$ for all players, and the conditions that make h the optimal choice in the first period become:

$$\Pr(\alpha = 1|h) - \Pr(\alpha = 1|i) \geq K - 1 \qquad \text{for } \theta_A = 1 \tag{6}$$

$$\frac{1}{K}(\Pr(\alpha = 1|h) - \Pr(\alpha = 1|i)) \geq K - 1 \qquad \text{for } \theta_A = K \tag{7}$$

$$\Pr(\alpha = 0|h) - \Pr(\alpha = 0|i) \geq K - 1 \qquad \text{for } \theta_B = 1 \tag{8}$$

$$\frac{1}{K}(\Pr(\alpha = 0|h) - \Pr(\alpha = 0|i)) \geq K - 1 \qquad \text{for } \theta_B = K. \tag{9}$$

Comparing the left-hand side of these equations, using assumption 5 (monotonicity of employer's strategies), it follows that

$$\Pr(\alpha|h) - \Pr(\alpha|i) \geq \frac{1}{K}(\Pr(\alpha|h) - \Pr(\alpha|i)).$$

In other words, there cannot be an equilibrium in which a worker with a higher cost of effort is more productive than a worker with a low cost of effort.

Let us try equation (4) as a candidate equilibrium strategy for the non-discriminatory employers. Equation (4) establishes that the employer promotes the worker displaying the higher productivity in the first period, while a coin is tossed when productivity is the same. Proposition 1 considers a situation in which both types of population A choose the same action of the corresponding type of population B. There are three possible situations: a) all the workers choose h, b) all the workers choose i, c) $e^1_{\theta=K} = i$ and $e^1_{\theta=1} = h$. Given that the distribution of types within populations is the same and that employer's beliefs about the distribution of workers' type are correct, in a) and b) the employer would be certainly indifferent. In c) an employer facing i, i or h, h is indifferent, while if h, i or i, h are observed, it is optimal to promote the worker who supplied the higher productivity, i.e. the "good" type. Hence the strategy is actually a best reply.

If there are non-discriminatory employers only, the equilibrium strategy (4) implies that for all workers

$$\Pr(\alpha = 1|h) - \Pr(\alpha = 1|i) = 0.5.$$

On the other hand, if there are also discriminatory employers:

$$\Pr(\alpha = 1|h) - \Pr(\alpha = 1|i) = 0.5 \Pr(\theta_F = 0). \tag{10}$$

Note that the incentive to supply h is the same for both populations even when there are discriminatory employers. This is intuitive, because the assumption that discriminatory employers always promote A makes the incentive to exert effort h proportional to the percentage of non-discriminatory employers. In the limit situation where there are only discriminatory employers, promotion stops being an incentive device for both populations, because the As are sure to be

promoted, while Bs have no chance. This parallels the finding within unfair tournaments that both agents exert the same level of effort in equilibrium.

Substituting the equation (10) into equations (6)-(9) above, the conditions that make h the optimal choice can be rewritten

$$0.5 \Pr(\theta_F = 0) > K - 1 \text{ for workers } \theta_A, \theta_B = 1,$$

$$\frac{1}{K} 0.5 \Pr(\theta_F = 0) > K - 1 \text{ for worker } \theta_A, \theta_B = K.$$

The presence of a strictly positive fraction of non-discriminatory employers is necessary for promotions to work as an incentive device. According to the parameter K different combinations of effort are observed in the first period, all representing a PBNE with the characteristics 1) and 2) of proposition 1 and consistent with the candidate equilibrium strategy for the employer proposed in (4).

Regardless of the value of K, the fraction of workers A who are promoted is always $1 - 0.5 \Pr(\theta_F = 0)$ and when $\Pr(\theta_F = 0) = 1$, i.e. when discriminatory tastes disappear, promotions are balanced across populations. In the second period, and in both populations, "good" workers who are promoted supply h, "bad" workers who are promoted as well as "good" workers who are not promoted supply i, while "bad" workers who are not promoted supply l.

The PBNE described in Proposition 1 are not unique. For instance, there are other PBNE associated with strategies of the employers different from (4). Outcomes of these equilibria can differ from those characterized above. However, these other PBNE share the feature that, for every equilibrium with more than $1 - 0.5 \Pr(\theta_F = 0)$ workers A promoted, there exists another equilibrium in which $0.5 \Pr(\theta_F = 0)$ workers B are promoted.[21]

What changes if $\mu^B(\theta_F) = \Pr(\theta_F)$ no longer holds and in particular when workers B overestimate the percentage of discriminatory employers? Beliefs about opponents' behavioral rules are still correct for all players by assumption. Hence, the equilibria that arise are still PBNE. Assuming that $\mu^B(\theta_F) \neq \Pr(\theta_F)$ while assumption 6a still holds means that beliefs do not satisfy the Common Prior assumption anymore. As already mentioned, in this case beliefs about opponents' type may be not only incorrect but also contradicted by the evidence. However, Proposition 2 refers only to PBNE that are also SCE, i.e. where beliefs are not contradicted by the evidence.

Proposition 2 *If a Self-Confirming and Perfect Bayes-Nash Equilibrium exists when minority workers overestimate the percentage of discriminatory employers, i.e. when $\mu^B(\theta_F = d) > \Pr(\theta_F = d)$, it must be characterized by*
1) all workers A supplying h and all workers B supplying i
2) only workers A being promoted.

---

[21] Characteristics of equilibria different from those proposed in Proposition 1 have been analyzed by means of simulations.

Conditions under which unequal outcomes arise in equilibrium.

For such an equilibrium to exist, conditions for the incentive of workers A to supply h do not change with respect to (6) and (7). As far as workers B are concerned, instead, conditions (8) and (9) become:

$$\mu^B\left(\alpha = 0|h\right) - \mu^B\left(\alpha = 0|i\right) \quad > \quad K - 1 \quad \text{for } \theta = 1, \tag{11}$$

$$\frac{1}{K}\left(\mu^B\left(\alpha = 0|h\right) - \mu^B\left(\alpha = 0|i\right)\right) \quad > \quad K - 1 \quad \text{for } \theta = K. \tag{12}$$

The same strategy as in (4) for non-discriminatory employers, together with the assumption that discriminatory employers promote only workers A, implies that the gain in the probability of being promoted playing h instead of i is

$$\Pr(\alpha = 1|h) - \Pr(\alpha = 1|i) = 0.5\Pr\left(\theta_F = 0\right), \tag{13}$$

$$\mu^B(\alpha = 0|h) - \mu^B(\alpha = 0|i) = 0.5\mu^B\left(\theta_F = 0\right), \tag{14}$$

for workers A and B, respectively.

Combining (8), (9), (11), (12), (13) and (2) the conditions that make h the optimal choice can be rewritten:

$$0.5\Pr\left(\theta_F = 0\right) \quad \geq \quad K - 1 \quad \text{for } \theta_A = 1,$$

$$\frac{1}{K}0.5\Pr\left(\theta_F = 0\right) \quad \geq \quad K - 1 \quad \text{for } \theta_A = K,$$

$$0.5\mu^B\left(\theta_F = 0\right) \quad \geq \quad K - 1 \quad \text{for } \theta_B = 1,$$

$$\frac{1}{K}0.5\mu^B\left(\theta_F = 0\right) \quad \geq \quad K - 1 \quad \text{for } \theta_B = K.$$

According to the values taken by the parameter K, different situations emerge. The condition

$$\frac{1 + \sqrt{1 + 2\Pr\left(\theta_F = 0\right)}}{2} \geq K \geq 1 + 0.5\mu^B\left(\theta_F = 0\right) \tag{15}$$

ensures that all workers B supply i and all workers A supply h in line with the first part of the proposition.[22] The fact that workers B overestimate the percentage of discriminatory employers is a necessary condition for these inequalities to hold.[23]

Since all the workers of population A supply h and all the workers of population B supply i, the employers' strategy (4) implies that only workers A are promoted (second part of the proposition). Such a strategy is in turn a best reply for the employers. The expected productivity in the second period being the same regardless of which worker is promoted, the employer is actually indifferent.

---

[22] For instance, with $K = 1.2$, the inequalities are satisfied if at the same time $\mu^B\left(\theta_F = 0\right) < 0.4$ and $\Pr\left(\theta_F = 0\right) \geq 0.5$.

[23] The result 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* equal workers.

Empiricism

Workers A and employers have correct beliefs about other players' type-strategy profiles. Hence, the objective distribution of observables coincides with the subjective distributions implied by their beliefs.

Wrong beliefs of workers B concern the employers' type only. Assumption 6b implies that their expected distributions of productivity (and therefore wages) within populations in the first period are correct. Associated with the observed outcome, "worker A supply h - worker B supply i," their correct beliefs about employers' strategies are associated with no worker B promoted even though their beliefs about employers' type are wrong. Finally, also the expected distribution of wages within populations in the second period is correct.

Uniqueness of unequal outcomes when $\mu^B(\theta_F = d) > \Pr(\theta_F = d)$

Observe first that uniqueness refers to the vector of observables (productivity and promotions) and not to the array of strategies and beliefs that characterize an equilibrium. In other words, there can be many observationally equivalent equilibria, i.e. many arrays of strategies and beliefs that generate the same vector of observables. Moreover, the importance of the uniqueness of such equilibria does not go beyond the goal of addressing a possible question of the reader, who could otherwise reasonably wonder whether there are other equilibria and what characteristics they have. Needless to say, uniqueness relies upon the whole set of assumptions that have been made, not merely upon $\mu^B(\theta_F = 0) < \Pr(\theta_F = 0)$.

In order to show the uniqueness of the vector of observables described in Proposition 2, the first step is to delete all the combinations of productivity levels associated with the combination of strategies that have no chance of being chosen.[24] Among the various combinations of productivity levels that can be observed, it can be shown that $\pi_A^1 = h, \pi_B^1 = i$ is the only one compatible with an employers' best response that does not falsify minority workers' wrong beliefs. A necessary condition for the wrong beliefs not to be contradicted is that also the non-discriminatory employers promote worker A after observing $\pi_A^1 = h, \pi_B^1 = i$.

The hypotheses behind Propositions 1 and 2 differ only because of the expectations of workers B. In Proposition 1 their 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 unequal outcomes with only workers A promoted.[25] Furthermore, nothing changes from the theoretical point of view

---

[24] For instance, given assumption 5, it cannot happen that a "bad" type of worker exerts a strictly higher effort than a "good" type of the same population in the first period. Furthermore, $\mu^B(\theta_F = 0) < \Pr(\theta_F = 0)$ implies that for a worker B it cannot be optimal to supply a strictly higher effort than the corresponding type of worker A, given that they share the same beliefs about employers' strategies (assumption 7).

[25] Observing only workers A promoted is certainly a knife-edge result. This is due to the strong assumptions made throughout the chapter. Having more than two types of workers, for instance, makes it possible to observe equilibria in which the fraction of workers A promoted is greater than that of the corresponding PBNE with Common Prior, but lower than one.

when it is assumed that $\Pr(\theta_F = d) = 0$ (absence of discriminatory employers), meaning that workers' expectations can be a "stand alone" source of unequal outcomes.

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 with exerting high effort can the initial beliefs be contradicted, but this cannot happen because of a "free-riding" problem.

Comparing results in Proposition 1 with what happens in Proposition 2, it turns out that workers B are worse off while workers A are better off, because of the change in the probability of being promoted that become more favorable for the latter. Also employers are worse off. Being proportional to workers' productivity, profits are lower in the first period given that workers B supply i rather than h while profits do not change in the second period.

4.4 Policy implications

Trial work periods can be an effective policy tool to break down statistical discrimination outcomes, i.e. a situation where employers' wrong beliefs are self-confirming. On the contrary, the equilibria described in Proposition 2 are robust to trial work periods, for the simple reason that trial work periods affect employers' expectations rather than workers' expectations. As long as minority workers think they are discriminated against, during the trial work period they will display a lower productivity. At the end of the first period of the game, which can be regarded as a long trial work period, employers observe workers A supplying a higher output and promote them even though there is no bias, either statistical or driven by tastes, against the minority.

Quotas can also be implemented to correct unequal outcomes. Although they effectively increase the number of minority workers promoted, they do so without affecting the mechanism that generates unequal outcomes in the equilibria of Proposition 2. The simplest way to implement quotas is to say 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 in every constituent game, such a result can be obtained by imposing a lottery on 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. Paradoxically, after the introduction of quotas workers B are less likely to exert effort h in the first period. In fact, conditions for h being an

optimal choice become:[26]

$$(1 - q)0.5\mu^{B}(\theta_F = 0) \geq (K - 1) \text{ for } \theta_B = 1,$$
$$(1 - q)\frac{1}{K}0.5\mu^{B}(\theta_F = 0) \geq (K - 1) \text{ for } \theta_B = K.$$

The same happens for workers A:

$$(1 - q)0.5\Pr(\theta_F = 0) \geq (K - 1) \text{ for } \theta_A = 1,$$
$$(1 - q)\frac{1}{K}0.5\Pr(\theta_F = 0) \geq (K - 1) \text{ for } \theta_A = K.$$

If for workers A it is still optimal to supply h, nothing changes with respect to Proposition 2, except that now all players correctly expect that q minority workers will be promoted. For minority workers to realize that they are overestimating the percentage of discriminatory employers, the only way is to impose a q "big enough" to make (one or both types of) workers A choose i instead of h. At that point, the number of minority workers who are promoted will be strictly greater than q, thus contradicting their beliefs.[27] It is worth noting that the price for such a result, when achievable, is that both majority workers and employers are strictly worse off with the introduction of quotas, which probably makes it not rather difficult to implement. A similar trade-off between equity and efficiency associated with affirmative action programs can also be found within uneven tournaments. Experimental evidence, however, does not provide support for such a trade-off.[28]

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

> "a 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. Many devices can perform this task, like a subsidy to minority workers who exert effort h, or a free insurance that pays back the money equivalent of the utility loss suffered by minority workers who supplied a high effort without being promoted, and so on. However, these devices do not seem to have an intuitive counterpart on the field. The

---

[26] Notice that the LHS is negatively correlated with q.

[27] It is problematic to provide a sensible translation of "big enough," since this threshold, depending on many factors, varies considerably. For instance, in a situation without discriminatory tastes and $K = 1.2$, i.e. around the middle of the range that makes it optimal for both types of worker A to exert h, even imposing balanced promotions, i.e. $q = 0.5$, is not enough to break down the mechanism behind unequal outcomes based on workers' expectations.

[28] See Schotter and Weigelt (1992) and Corns and Schotter (1999).

bottom line is that, in line with Coate and Loury (1993), the best way to correct unequal outcomes is to modify the expectations of minorities.[29] Policy tools which do not change the expectations of minorities are either ineffective or very difficult to implement. Filippin (2003b) presents a laboratory experiment that provides empirical evidence supporting the importance of workers' expectations as a source of unequal outcomes.

5   Conclusions

The goal of this paper is to set up a model where the preferences and beliefs of both sides of the labor market matter. A framework is obtained where most of the contributions to the discrimination literature can be nested. Moreover, the role of workers' expectations, almost neglected in the literature, can be analyzed. A game of incomplete information is presented, showing that *ex ante* identical groups of workers may be characterized by unequal outcomes in equilibrium because of their different beliefs. The importance of workers' expectations can be appreciated by comparing the distribution of promotions that arises when minority workers overestimate the percentage of discriminatory employers with a situation in which such beliefs are correct *ceteris paribus*. With the purpose of testing workers' expectations as a "stand alone" mechanism, the comparison is made by imposing appropriate assumptions that rule out other possible sources of unequal outcomes.

Unequal outcomes may arise due to workers' expectations. In this situation what happens is that wrong beliefs of being discriminated against are self-confirming. Minority groups who expect of being discriminated against supply less effort on average, because of a lower expected return. This induces a lower percentage of promotions within minority workers, even though employers do not discriminate against them either directly or statistically. Nevertheless, unbalanced promotions are consistent with their beliefs that there are employers with discriminatory tastes. This mechanism implies that workers' expectations of being discriminated against are important in reducing the effectiveness of promotion as an incentive device and they can contribute to the puzzling long-run persistence of cumulative discrimination.

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. Moreover, trial work periods, which can break down statistical discrimination outcomes, are not an effective policy tool as long as workers have expectations of employers' discriminatory tastes. Furthermore, wrong beliefs of minority workers are unlikely to be modified by the introduction of quotas. The game suggests that the best way to get rid of unequal outcomes driven by workers' expectations is by using beliefs themselves as a target.

---

[29]For instance, Gay Pride can also be thought of as a device that reduces the sexual minorities' expectations of being discriminated against in the labor market.

References

[1] AIGNER, D.J. AND G.G. CAIN (1977): "Statistical Theories of Discrimination in Labor Markets," *Industrial and Labor Relations Review*, 30:2, 175-87.

[2] 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.

[3] BATTIGALLI, P. (1987): *Comportamento Razionale ed Equilibrio nei Giochi e nelle Situazioni Sociali*, unpublished dissertation, Universita' Bocconi, Milano.

[4] 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. Dordecht: Kluwer Academic Publishers, 97-124.

[5] BAYE, M.R., D. KOVENOCK AND C.G. DE VRIERS (1996): "The All-Pay Auction with Complete Information," *Economic Theory*, 8, 291-305.

[6] BECKER, G.S. (1957): *The Economics of Discrimination*. Chicago: University of Chicago Press.

[7] ——— (1985): "Human Capital, Effort and the Sexual Division of Labor," *Journal of Labor Economics*, 3, S33-S58.

[8] BLAU, F.D., M.A. FERBER AND A. E. WINKLER (2002): *The economics of Women, Men and Work*. Upper Saddle River, NJ: Prentice Hall.

[9] 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.

[10] BREEN R., AND C. GARCIA-PENALOSA (2002): "Bayesian Learning and Gender Segregation," *Journal of Labor Economics*, 20:4, 899-922.

[11] 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.

[12] COATE, S. AND G.C. LOURY (1993): "Will Affirmative-Action Policies Eliminate Negative Stereotypes?," *American Economic Review*, 83:5, 1220-40.

[13] 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.

[14] DEKEL, E., D. FUDENBERG AND D.K. LEVINE (2003): "Learning to Play Bayesian Games," *Games and Economic Behavior,* forthcoming.

[15] FERBER, M.A. AND H.M. LOWRY (1976): "The Sex Differential in Earnings: A Reappraisal," *Industrial and Labor Relations Review*, 29, 377-87.

[16] FILIPPIN, A. (2003): "A Generalization of the Self-Confirming Equilibrium Concept," mimeo.

[17] ———— (2003b): "Discrimination and Workers' Expectations: Experimental Evidence," mimeo.

[18] FILIPPIN, A. AND A. ICHINO (2003): "Gender Wag Gap in Expectations and Realizations," mimeo.

[19] FRYER, R. (2003): "A Dynamic Theory of Statistical Discrimination," mimeo.

[20] FUDENBERG, D. AND D.K. LEVINE (1993): "Self-Confirming Equilibrium," *Econometrica*, 61:3, 523-545.

[21] ———— (1998): *The Theory of Learning in Games.* Cambridge and London: MIT Press.

[22] LAZEAR, E. AND S. ROSEN (1981): "Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy*, 89, 841-864.

[23] MINCER, J. AND S. POLACHECK (1974): "Family investments in Human Capital: Earnings of Women," *Journal of Political Economy*, 82:2, 76-108.

[24] MORO, A., AND P. NORMAN (2003): "A General Equilibrium Model of Statistical Discrimination," *Journal of Economic Theory,* forthcoming.

[25] MYRDAL, G. (1944): *An American dilemma.* New York: Harpher & Brothers.

[26] O'KEEFFE, M., W. K. VISCUSI, AND R. J. ZECKHAUSER (1984): "Economic Contests: Comparative Reward Schemes," *Journal of Labor Economics*, 2:1, 27-56.

[27] PHELPS, E.S. (1972): "The Statistical Theory of Racism and Sexism," *American Economic Review*, 62:4, 659-661.

[28] SCHOTTER, A. AND K. WEIGELT (1992): "Asymmetric Tournaments, Equal Opportunity Laws, and Affirmative Action: Some Experimental Results," *Quarterly Journal of Economics*, 107, 511-39.

[29] SPENCE, M.A. (1973): "Job Market Signaling," *Quarterly Journal of Economics*, 87:3, 355-374.

IZA Discussion Papers

No.	Author(s)	Title	Area	Date
809	S. Commander M. Kangasniemi L. A. Winters	The Brain Drain: Curse or Boon?	1	06/03
810	J. H. Abbring G. J. van den Berg	A Simple Procedure for the Evaluation of Treatment Effects on Duration Variables	6	06/03
811	M. Corak W.-H. Chen	Firms, Industries, and Unemployment Insurance: An Analysis Using Employer-Employee Data	1	06/03
812	J. T. Addison T. Schank C. Schnabel J. Wagner	German Works Councils in the Production Process	3	06/03
813	E. P. Lazear	Firm-Specific Human Capital: A Skill-Weights Approach	5	06/03
814	G. Ridder G. J. van den Berg	Measuring Labor Market Frictions: A Cross-Country Comparison	6	07/03
815	A. Aakvik K. G. Salvanes K. Vaage	Measuring Heterogeneity in the Returns to Education in Norway Using Educational Reforms	6	07/03
816	T. T. Herbertsson J. M. Orszag	The Early Retirement Burden: Assessing the Costs of the Continued Prevalence of Early Retirement in OECD Countries	3	07/03
817	T. M. Andersen T. T. Herbertsson	Measuring Globalization	2	07/03
818	J. Pencavel	The Surprising Retreat of Union Britain	3	07/03
819	M. Beine F. Docquier H. Rapoport	Brain Drain and LDCs' Growth: Winners and Losers	1	07/03
820	C. M. Cornwell K. H. Lee D. B. Mustard	The Effects of Merit-Based Financial Aid on Course Enrollment, Withdrawal and Completion in College	6	07/03
821	P. Carneiro J. J. Heckman	Human Capital Policy	6	07/03
822	D. Weichselbaumer R. Winter-Ebmer	The Effects of Competition and Equal Treatment Laws on the Gender Wage Differential	6	07/03
823	A. Filippin	Discrimination and Workers' Expectations	5	07/03

An updated list of IZA Discussion Papers is available on the center's homepage www.iza.org.