Discrimination and Workers' Expectations

Antonio Filippin
European University Institute
antonio.filippin@iue.it

February 13, 2002

Abstract

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.

Keywords: Discrimination, Workers' Expectations, Self-Confirming Beliefs.

JEL classification codes: J71, J15, J24, D84, C79.

1 Introduction

Despite the several contributions in the literature, a widely shared explanation for the long run persistence of observed discrimination in the labor markets is still not apparent. Many theories provide a wide range of possible explanations for the presence of discrimination. However, these theories fail to be fully convincing as far as their long run implications are concerned. Among the most important contributions, there are the so called "discriminatory tastes" and "statistical discrimination" models.[1] However, as Cain (1986) points out

> "the neoclassical theory of discrimination is almost entirely a demand-side theory. The supply-side of the labor market is effectively neutralized by the assumption that minority and majority groups of workers are equally productive (or have equal productivity capacity) and have equal tastes for work. The demand side may be characterized ... by 'exact' versus 'stochastic' models."

[1] More details about these models, as well as the the human capital theory and the feedback effect approach are contained in section 2.3, where these contribution are nested into the model presented in this paper.

The goal of this paper is to fill this gap, setting 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 the focus of this paper.

The model is formalized as a two-stage game of incomplete information where populations of workers and employers are engaged. In every stage game two workers, one of which is a minority worker, and one employer are randomly matched. The employer promotes one of the two workers after having observed their output, which is a function of effort and tastes for work, both unobservable. Promotions also depend on employer's type, unknown to the workers, which captures the possible disutility of promoting a minority worker.

The importance of workers' expectations can be appreciated comparing the equilibrium outcome in terms of promotions that arises when minority workers overestimate the percentage of employers characterized by tastes for discrimination with a situation in which beliefs are correct *ceteris paribus*.[2] Even in a labor market where discriminatory tastes have disappeared, statistical discrimination is absent and all the other sources of heterogeneity such as distribution of ability among workers, etc. have been neutralized, unequal outcomes may arise due to workers' expectations. In this circumstance what happens is that in equilibrium wrong beliefs to be discriminated against are self-confirming. Minority groups who are (or expect to be) discriminated against supply less effort on average, because of a lower expected return. This induces a lower percentage of promotions within minority workers, which in turn convinces them that there are employers with discriminatory tastes.

The structure of the paper is the following. After some definitions are outlined in Section 1.1, the stage game of the model, i.e. the game after the players have already been matched, is presented (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 2.3. Section 3 concentrates on the analysis of the equilibria of the model. Section 4 concludes.

1.1 Definitions

Before going on with the presentation of the model, it is useful to clarify some concepts.

To avoid confusion, in what follows **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 tastes for work. Therefore, a more productive worker is simply characterized by a higher output.

[2] As explained in Section 3, correct beliefs characterize Bayes-Nash Equilibria, wrong beliefs characterize Self-Confirming Equilibria.

Having specified what productivity means in this paper, the following question is: what is discrimination? In the literature many different and sometimes contradicting definitions have been used. In fact, discrimination in the labor market 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 *tastes for work*. Sometimes 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, referring in part to Ferber and Lowry (1976), 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.[3] On the other hand, also a more comprehensive definition seems to be necessary. 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 effects of discrimination on workers' behavior. If minority workers are less productive for example because they have changed their behavior reacting to direct discrimination, the different achievements should not be viewed as equal treatment, even if there is no more direct discrimination. Therefore, it is useful also the more comprehensive concept of **cumulative 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 average achievements are observed either for groups of workers which are on average equally productive (direct group discrimination) or for groups of workers which on average are ex ante equal (cumulative group discrimination). The latter happens when an individual is judged also 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, on which this paper focuses. 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 where populations of workers and employers are engaged. The two population 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

[3] In Ferber and Lowry (1976) the definition of direct discrimination refers to "equally qualified" workers.

called minority worker, identified by superscript B. Productivity, equal to effort $\pi = e$, is assumed to be measurable and to be entirely paid to workers.[4] Employers are denoted by superscript F.

The following section focuses on the stage game, i.e. what happens after the players have already 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 stage game

In every stage game two workers, one of which is a "minority" worker, and one employer are drawn from their populations and play. In the first period both workers choose one out of three possible levels of effort $e_1^A, e_1^B \in E = \{l, i, h\}$, where l stands for "low", i stands for "intermediate" and h stands for "high". The employer observes workers' productivity in the first period and decide which of the two workers to promote. After having observed employer's decision, workers choose a level of effort for the second period $e_2^A, e_2^B \in E$.

The stage 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 with effort. The game is also characterized by incomplete information, because every player knows his or her type only.

2.1.1 Incomplete information

$m^A \in M$ and $m^B \in M$, summarize the type of majority and minority workers, respectively. Workers' type represents their tastes for work, formalized as a weight attached to the disutility of effort (see equation 1 and 2 below). The lower the parameter, the lower the cost of effort.

$m^F \in M^F = \{d, 0\}$ represents tastes for discrimination of employers, which are restricted to be binary: $m^F = d > 0$ if the employer suffer a disutility when the minority worker is promoted, and $m^F = 0$ if the employer is indifferent about the observable characteristic which distinguishes the workers.

Assumption 1: *Every player knows her own type only.* Therefore, a minority worker knows her own tastes for work m^B, while the type m^A of the majority worker and the tastes for discrimination m^F of the employer are unknown.[5]

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 the type m of the worker.

[4] The assumption that $\pi = e$ implies that ability does not matter in this model. A more general version in which ability is not constant among workers and its distribution is the same within populations of workers turns out to be much more complicate without being more insightful (see also footnote 5).

[5] Of course, the same holds mutatis mutandis for player B and F.

Workers' lifetime utility is

$$U^{m,A} = w_1^A - \left(e_1^A\right)^2 + \alpha\left(w_2^A - \frac{m^A}{K}\left(e_2^A\right)^2\right) + (1-\alpha)\left(w_2^A - m^A\left(e_2^A\right)^2\right) \quad (1)$$

$$U^{m,B} = w_1^B - \left(e_1^B\right)^2 + (1-\alpha)\left(w_2^B - \frac{m^B}{K}\left(e_2^B\right)^2\right) + \alpha\left(w_2^B - m^A\left(e_2^B\right)^2\right) \quad (2)$$

where:

w_t^P is the wage in period t for the worker belonging to population P.

e_t^P is effort in period t for the worker belonging to population P.

$\alpha = 1$ 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.

Since labor market is assumed to be competitive, productivity is entirely paid to workers. Moreover, productivity coincide with effort. As a consequence, the optimal level of effort in the second period turns out to be

$$\begin{aligned} e_2^{A*}|\,(\alpha=1) &= \frac{K}{2m^A} > e_2^{A*}|\,(\alpha=0) = \frac{1}{2m^A} \\ e_2^{B*}|\,(\alpha=0) &= \frac{K}{2m^B} > e_2^{B*}|\,(\alpha=1) = \frac{1}{2m^B}. \end{aligned}$$

Not surprisingly, effort will be higher if the worker is promoted.[6]

Assumption 2: *The set of types is $M_1 = \{1\}$ in the first period and $M_2 = \{1, K\}$ in the second.* Tastes for work are assumed to matter in the second period only. In the first period tastes for work are the same for all workers.[7] There are only two possible types of worker. This assumption also implies that in the second period a bad type if promoted supplies the same effort of a good type who is not promoted. Therefore, in the second period there are only three conceivable levels of optimal effort.

Assumption 3: *The set of levels of effort E is* $\left\{l = \frac{1}{2K}, i = \frac{1}{2}, h = \frac{K}{2}\right\}$. The three levels of effort l, i, h coincide with the optimal level in the second period for a bad type not promoted, a bad type promoted or a good type not promoted, a good type promoted, respectively.

As far as the employer is concerned, the utility function contains both profits and a parameter summarizing the disutility associated to the promotion of

[6] Notice that the same results would be obtained if K was some know-how accessible to the promoted worker only, which multiplies her effort in the second period:

$$U^{m,A} = w_1^A - \left(e_1^A\right)^2 + \alpha K e_2^A + (1-\alpha)\, w_2^A - m^A\left(e_2^A\right)^2$$

[7] This assumption is important. Allowing different tastes for work in the first period would *de facto* resolve 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 3. It would be possible to relax this assumption if at the same time ability was not restricted to be costant for all workers. A more plausible model would be obtained providing the same insights at the price of increased complication.

the worker B. This means that only worker B faces the risk to be discriminated against, because of the observable characteristic that without entering productivity differentiates her from worker 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

$$U^F = (u-1)\left(\pi_1^A + \pi_1^B + \pi_2^A + \pi_2^B\right) - (1-\alpha)m^F \tag{3}$$

where $u > 1$ is a known and constant mark up on workers' productivity due to the entrepreneurial activity. Therefore, $u - 1$ represents profits as a function of workers productivity. In order to maximize profits, the employer needs to maximize worker's productivity. In other words, the employer has an incentive to promote the more productive worker.[8] The term $(1-\alpha)\, m^F$ represents the disutility associated to a promotion of a minority worker. When $m^F = 0$ 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 $m^F = d > 0$ the employer is characterized by discriminatory tastes, and the minority worker is less likely to be promoted.

2.1.3 Strategies

Workers move twice, the second time after the decision about promotion has been taken by the employer, choosing simultaneously a level of effort. The strategy of a worker s^A, s^B is therefore a triple containing a level of effort for the first period, a level of effort for the second period if promoted and a level of effort for the second period if not promoted. For both populations of workers effort can take the same three values only: $e_1, e_2 \in E = \{h, i, l\}$, where $h > i > l > 0$ stand for "high", "intermediate" and "low", respectively. Therefore, the strategy set is the same for both workers, i.e. $S^A = S^B$. Moreover, the strategy set is assumed to be type independent, meaning that every type of worker faces the same possible choices.

The employer observes each worker's productivity, but in general it is not possible to recover the worker's type from her productivity. After having observed the output of the two workers, the employer select one of the two workers and promotes her in a more rewarding position. The set of feasible actions for the employer, regardless of her type, is simply $\alpha = \{0, 1\}$, where $\alpha = 1$ stands for "promote worker A" and $\alpha = 0$ stands for "promote worker B". As far as the employer is concerned, strategies s^F are therefore a vector that specifies the decision about promotion for every possible couple of productivity levels.

Assumption 4: *disutility* $d \to \infty$. For the sake of simplicity disutility d is assumed to be high enough to make a discriminatory employer always promoting worker A.

[8] 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. Therefore, nothing would change in practice, because also the supervisor has the incentive to promote the more productive worker in order to maximize her bonus.

As already mentioned in section 2.1, a non discriminatory employer does not care about the observable characteristics which distinguishes workers. Hence, employer will simply promote the worker that is expected to be more productive in the second period. If workers are expected to be equally productive in the second period, the employer breaks the tie in favor of the worker who displayed the highest productivity in the first period. If also productivity in the first period was the same, the employer tosses a coin.

To complete the description of this stage game also the players' beliefs 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 stage game described in section 2.1 is inserted in a wider game, called population game, described in this section, which specifies how players are matched and what information they can access. The description of the information structure allows to define players' beliefs.

There are three populations, one of employers and two of workers. As already said for the stage game, the two populations of workers differ because of an observable characteristic (e.g. gender, race, etc.) that does not enter workers' productivity.

Assumption 5: *The distribution of types within the two populations of workers is identical.* This assumption rules out the possibility that differentials in promotions arise because one population of workers has higher tastes for work than the other.[9]

2.2.1 Matching

Each of the three populations $P = \{A, B, F\}$ is composed by an infinite number of identical cohorts, and every cohort is composed by a continuum of players. The assumption of a continuum of players is necessary to invoke the law of large numbers which ensures a deterministic combination of types after the random matching.[10] The three populations play an infinitely repeated game, while each cohort plays only one "round", i.e. a two-stage game. At every stage only one cohort from population A, one cohort from population B, and one cohort from population F play. Each of the "active" players, i.e. the players of the only cohort engaged in the game, of population A is randomly matched with one of the "active" players of population B and one of the "active" players of population F.

[9] As we will see in section 2.3, this is not the case for the human capital approach.

[10] As it will be explained later in section 3, in this way, the objective and the subjective distributions of observables can coincide (empiricism). This is one of the requirements of the conjectural (or self-confirming) equilibrium.

2.2.2 Information structure

Although all cohorts within each population are equal, identifying different cohorts is far from being irrelevant. On the contrary, cohorts are very important to characterize the information structure. In particular, the division of the population into different cohorts captures a very important feature: the game is repeated infinitely among populations, but each player participate in just one round. This has the very important implication that individual outcomes are a useless source of information. On the contrary, aggregate outcomes of previous rounds can be informative.[11]

Aggregate outcomes consist of an array of distributions of observables $\hat{\sigma} = (\hat{w}_1^A, \hat{w}_1^B, \hat{\alpha}, \hat{w}_2^A, \hat{w}_2^B)$, where $\hat{w}^A, \hat{w}^B$ are the distributions of wages (equal to productivity) within each population of workers in both periods. $\hat{\alpha}$ is the percentage of workers A promoted and therefore summarizes the distribution of promotions across workers A and B.

Assumption 6: *At the beginning of each round, all players know the distributions of observables which arose in the previous round.*

2.2.3 Beliefs

Beliefs of a player are a probability measure over the unknown component of the type-strategy set $M \times S = M^A \times M^B \times M^F \times S^A \times S^B \times S^F$. Given that every player is supposed to know her own type and the strategy she chooses only, the unknown component of $M \times S$ turns out to be the set of type-strategy profiles of all the other players, both the opponents and the other players of her own population. Beliefs of a worker of population A when her type is m are defined[12]

$$\mu^{m,A} \in \ \Delta\left(M \times S\right). \tag{4}$$

Assumption 7: *Beliefs exclude the possibility that opponents correlate their play.* In other words, it holds that:

$$\mu^{m,A}(\mathsf{m}^A, \mathsf{s}^A, \mathsf{m}^B, \mathsf{s}^B, \mathsf{m}^F, \mathsf{s}^F) = \mu^{m,A}(\mathsf{m}^A, \mathsf{s}^A)\mu^{m,A}(\mathsf{m}^B, \mathsf{s}^B)\mu^{m,A}(\mathsf{m}^F, \mathsf{s}^F).$$

In the stage game, every player knows her own type and strategy and therefore the appropriate marginal distribution is used. For instance, the relevant beliefs of worker A are:

$$\mu^{m,A}(\mathsf{m}^B, \mathsf{s}^B, \mathsf{m}^F, \mathsf{s}^F) = \mu^{m,A}(\mathsf{m}^B, \mathsf{s}^B)\mu^{m,A}(\mathsf{m}^F, \mathsf{s}^F). \tag{5}$$

[11] Using aggregate outcomes instead of individual outcomes of some elements of previous cohorts also rules out the possibility that unequal outcomes arise in equilibrium as an inerhitance of past difference in fundamentals, as in Breen and Garcia-Penalosa (1998).

[12] When the type is used as an index, like in this case for beliefs, the notation $(\cdot)^{\mathsf{m},\mathsf{A}}$ is used instead of $(\cdot)^{\mathsf{m}^{\mathsf{A}}}$.

Something more can be said about employers' beliefs. Employers have the opportunity to update their beliefs about workers before the decision about promotion, after having observed their productivity in the first period. Employers' prior beliefs are a probability measure over workers' type-strategy profile $\mu^{m,F}\left(m^A, s^A, m^B, s^B\right)$. Such beliefs can be revised independently using Bayes rule when productivity is observed, given that worker A's productivity does not convey information about worker B and viceversa. Defining

$$\Pr\left(\pi_1^A\right) = \sum_{(m^A, s^A) \in (M^A \times S^A)} \mu^{m,F}(m^A, s^A : e_1^A = \pi_1^A) \tag{6}$$

the probability that an employer of type m assigns productivity level $\pi_1^A \in \Pi^A$ according to her prior beliefs, revised beliefs after the observation of a productivity level $\tilde{\pi}_1^A$ will be:

$$\mu^{m,F}\left(m^A, s^A | \tilde{\pi}_1^A\right) = \begin{cases} \frac{\mu^{m,F}(m^A, s^A)}{\Pr\{\tilde{\pi}_1^A\}} & \text{if } m^A, s^A : e_1^A = \tilde{\pi}_1^A \\ 0 & \text{if } m^A, s^A : e_1^A \neq \tilde{\pi}_1^A \end{cases} . \tag{7}$$

Although this paper does not deal with dynamics, I think it is necessary to justify at least intuitively where beliefs come from. To do this a minimum of "thought dynamic" is required. Beliefs of a player at time t are a function of the available information about aggregate outcomes arising from the previous period $\hat{\sigma}_{t-1}$.

It is worth noting that the same sequence of observables can lead to different beliefs. In other words, the same information about aggregate outcomes can be interpreted in different ways by every player. For example, workers can interpret the same distribution of promotions across populations A and B assigning different importance to the effort of the workers Vs. the propensity to discriminate of the employers. Of course, asymptotic empiricism requires that in equilibrium all the beliefs rule must generate subjective distributions of observables which coincide with the objective one. This means that in equilibrium different belief rules for identical workers can survive only if observationally equivalent.

Before characterizing the equilibria of the game presented so far, it is useful to review briefly some of the contributions to the discrimination literature, that can be nested in this model imposing appropriate assumptions.

2.3 Related literature

The model presented so far is flexible enough to capture, under appropriate assumptions, the main features of several important contributions to the discrimination literature.[13] One thing that must be taken into account is that all

[13] The main goal of the following survey of the literature concerning discrimination in labor markets is to shed some light on the area surrounding the model presented in this paper. The theories presented here have been chosen and outlined in such a way as to facilitate contrast and comparison with the model of workers' expectations described in section 2.1 and 2.2. Therefore, the choice of the contributions to be summarized is far from exhaustive,

these contributions focus on wages rather than promotions. However, this is not a big issue from the qualitative point of view because the main stylized facts of these theories can be replicated also focussing on promotions.

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

2.3.1 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 neoclassical 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.

The following are the assumptions that should be imposed into the model presented in section 2.1 and 2.2, in order to make it as close as possible to the discriminatory tastes approach:

1. the type set of the employer is a singleton $M^F = \{d\}$, with $d > 0$ which implies discriminatory tastes.

2. Beliefs are degenerate, assigning a probability equal to one to the true type-strategy profile of the opponents (absence of uncertainty). Expectations do not matter, and workers are perfect substitutes in the production function.

While in the Becker's model discrimination takes the form of different pay for equal work, in the game obtained imposing these assumptions discrimination takes the form promoting always 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.). However, 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 would suffer a negative utility. Alternatively, we should observe complete segregation. Since

concentrating only on the theoretical aspects of some competitive neoclassical models and institutional theories. Marxian and noncompetitive neoclassical models are among the theories not included in this survey. Also the relative weight assigned to various aspects of the summarized theories reflects primarily the necessity of the subsequent presentation, rather than some sort of consensus about what has been considered more important in the literature so far. Another reason for these choices is that many detailed surveys are already available (Blau (1981) and Cain (1986) among others).

both predictions are contradicted by empirical evidence, this model cannot be considered completely satisfactory.

2.3.2 Statistical discrimination

Within statistical discrimination models, group membership is assumed to convey information regarding individual characteristics, about which incomplete information is assumed. Ruling out the possibility that employers make systematic mistakes in predicting the average productivity of a group,[14] several models have been developed under the assumption of incomplete information, using different devices in order to explain the long run persistence of observed discrimination. Common to these models is the fact that, differently from Becker's, fundamentals are not relevant.

The most representative model of statistical discrimination has been proposed by Arrow (1973).[15] Employer's beliefs about the existence of different characteristics between (ex ante identical instead) groups turn out to be correct in equilibrium. 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 worker are ex ante equal but show a different productivity in equilibrium.

Statistical discrimination outcomes, as explained by Arrow, can be obtained in the game presented in this paper only accepting some changes to the structure of stage game.[16] Such an abrupt change to nest the Arrow's model is due to the fact that in Arrow's model the employer makes her decision using prior beliefs. Such a characteristic does not find a direct correspondence in this paper, where the decision about promotion is taken using updated beliefs.

1. At the beginning of the first period the assignment to different jobs is decided once and for all. The different jobs, a good one and a bad one, correspond to the job assigned in section 2.1 to the promoted and to the non promoted worker, respectively. Workers move twice, after the job assignment, choosing simultaneously a level of effort.

[14] Under perfect competition, employers who make systematic mistakes should not be more likely to survive in the long run than those with discriminatory tastes (Aigner and Cain, 1977).

[15] 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 pioneristic 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 relevant also in the context of the model of workers' expectation presented in this paper and are therefore explicitely addressed in what follows.

[16] One could certainly argue that strictly speaking Arrow's model cannot be nested in the model of section 2.1 and 2.2, given that it is necessary to define a different stage game.

2. Also strategies and utility functions presented in section 2 should be modified according to the previous point. Employers' strategies coincide now with feasible actions $\{\alpha = 1, \alpha = 0\}$, with $\alpha = 1$ standing now for "assign worker A to the good job and worker B to the bad job" and viceversa for $\alpha = 0$. Workers' strategies are now a quadruple containing two levels of effort if assigned to the good job and two levels of effort is assigned to the bad job.

3. While employer's payoff does not change, workers' payoff become

$$U^{m,A} = \left(w_1^A + w_2^A\right) - \left(e_1^A + e_2^A\right) - \left(1 - \alpha + \frac{\alpha}{K}\right) m^A \left(e_2^A\right) \tag{8}$$

$$U^{m,B} = \left(w_1^B + w_2^B\right) - \left(e_1^B + e_2^B\right) - \left(\frac{1-\alpha}{K} + \alpha\right) m^B \left(e_2^B\right) \tag{9}$$

4. Discriminatory tastes do not play any role, i.e. the set of employer types is a singleton $M^F = \{0\}$. Notice that workers' expectations do not matter any more, because when they move the uncertainty has already been resolved, given that the employer's action has been observed and there is just one type of employers.

5. Employers believe that minority workers are on average less productive than majority workers in the good job. Defining $\mu^{m,F}\left(\pi^A | \alpha = 1\right)$ 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

$$\mu^{m,F}\left(\pi^A | \alpha = 1\right) > \mu^{m,F}\left(\pi^B | \alpha = 1\right). \tag{10}$$

   Employer expect minority workers to be less productive, and therefore assign them to the bad job.

Assignment to the bad job causes the minority workers to supply a lower effort with the result that ex post they are effectively less productive, confirming employer's expectations, and leading also to wage differentials among majority and minority workers.

Statistical discrimination models have been heavily criticized by Cain (1986), who claim that

> "even if one did not have the faith that the competitive market would facilitate efficient signaling instruments and institutions ... this 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"

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 stage game, where updated beliefs are used to decide promotion and where the whole first period can be thought as a big trial work period.

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 other words, do workers' expectations and heterogeneity of workers' type matter? In section 3, which contains a characterization of the equilibria of the game presented in this paper a positive answer to such a question emerges.

2.3.3 The Human Capital Theory

Another branch of the literature, started by Mincer and Polacheck (1974), is the so-called human capital theory which analyzes the effects of voluntary choices of women in human capital investment. According to this theory, women decide to invest less than men in human capital 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 in the family (Becker, 1985). In this way, wage differentials, worse career path, and/or sex segregation could be explained by voluntary choices. If this was the case, the different achievements *would not be classified as discrimination*, given that they are observed for workers who are neither equally productive in equilibrium nor ex ante equal. A shortcoming of this theory is its limited field of application: different achievements based on anything other than gender cannot be explained by this approach.

The human capital theory, as outlined by Mincer and Polacheck can easily be nested into this model.

1. It is sufficient to assume that the cumulative distribution of types within population B first order stochastically dominates the cumulative distribution of A. Having an higher m means weighting more the disutility of effort: minority workers find more convenient to supply on average a lower effort and they will be therefore less productive.

A criticism that some economists moved to this approach (see the next subsection) is that the seemingly "voluntary" decision could itself be determined by discrimination, entering the definition of cumulative discrimination.

2.3.4 Feedback Effects

The boundaries of this approach are particularly uncertain,[17] and usually surveys concerning labor market discrimination use these models as a counterpart

[17] A large number of the so called "institutional" contributions may also fall into this category. Cain (1986) includes also the above-mentioned model by Arrow (1974) within this goup. 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

for other theories, without analyzing them separately. The reason for this 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 direct and/or cumulative 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 assumption is necessary to nest this approach into the game of section 2.1 and 2.2: the presence of workers' expectation is *per se* a way to formalize such feedback effects.

Once again, Cain (1986) is skeptical about the importance of models in which multiple equilibria, some of which suboptimal, are involved:

> "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. These criticism apply, however, to the particular mathematical formulation of the model and not to the reasonable view that economic outcomes are determined by multiple causes, some of which are noneconomic, and that feedback relationship are part of reality."

2.3.5 Workers' expectations

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

The above-mentioned articles and many others which are related do not take into account the potential role of workers' expectations in explaining observed discrimination, which is instead the focus of this paper. The only article I have found in the literature that explicitly uses the heterogeneity of workers' preferences and beliefs as an explanatory variable is that of Breen and Penalosa (1998), who give an explanation of the observed persistence of gender segregation in labor markets using a Bayesian learning approach. Workers, due to

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: "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). Many other contributions follow along the line of the vicious circle described by Myrdal: among the others: Piore (1970), Ferber and Lowry (1976).

imperfect information, do not know how much the probability of success in various occupations is affected by effort or by predetermined individual characteristics (such as gender). Agents try to learn about the role of effort by means of a Bayesian updating mechanism, and updated beliefs about importance of effort versus gender are transmitted to the next generation. In other words, the "prior" of a man (woman) is the belief received by his father (her mother), while the posterior is the belief updated accordingly 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 different equilibrium outcomes for the two groups, even once preferences are equal, meaning that past circumstances will continue to exert an influence and that expectations can be self-fulfilling.

The work of Breen and Penalosa focuses on the importance of the worker's expectations and is able to explain the persistence of discrimination. However, 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.

This contribution cannot be nested into the model presented in this paper, because it requires a dynamic framework. What distinguishes this model from the model of Breen and Penalosa, besides the static Vs. dynamic perspective, is essentially the information structure involving aggregate outcomes, and the possibility to analyze the joint effect of expectations and type heterogeneity for both sides of the labor market, which is the focus of section 3.

3 Analysis of the equilibria

Two different concepts are used to analyze the equilibria of the model: the Bayes-Nash Equilibrium and the Self-Confirming (or Conjectural) Equilibrium. The two concepts have in common the fact that

1) each player maximizes her utility given her beliefs, updated whenever possible, about every possible opponents' type-strategy profile.

What distinguishes the latter from the former is that

2a) in the Bayes Nash Equilibrium (BNE) beliefs are correct;

2b) in the Self-Confirming Equilibrium (SCE) beliefs are not necessarily correct, provided that they are not contradicted by the evidence.

3.1 Utility maximization given beliefs

a1) Employers.

Only workers' difference in productivity *after* the promotion matters, while the difference before does not. The reason is that the disutility m^F is associated to 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 cost m^F in order to decide which worker is convenient to promote.

Employers of type $m^F = 0$ 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. Therefore, they will always promote the worker they think will be more productive *after* the promotion, regardless of the population where she comes from.

If workers are of the same type the overall productivity and hence employer's utility after the promotion is the same regardless of the worker who is promoted. On the other hand, if workers' type is different the employer maximizes her utility promoting the worker characterized by higher tastes for work (i.e. lower m).

Defining $\mu^{0,F}\left(\pi_2^A|\pi_1^A\right)$ the subjective probability that a non discriminatory employer assigns to the productivity π_2 of worker A, having observed π_1^A, it follows that her best reply $BR^{0,F}\left(\pi|\mu^{0,F}\right)$ to the observed productivity $\pi = \left(\pi_1^A, \pi_1^B\right)$ given updated beliefs $\mu^{0,F}\left(\pi_2^A|\pi_1^A\right)$ will depend on the comparison of workers' expected productivity in the second period. Formally:

$$BR^{0,F}\left(\pi|\mu^{0,F}\right) = \{\alpha = 1\} \text{ if } \sum_{\pi_2^A \in \Pi_2^A} \mu^{0,F}\left(\pi_2^A|\pi_1^A\right) \pi_2^A > \sum_{\pi_2^B \in \Pi^B} \mu^{0,F}\left(\pi_2^B|\pi_1^B\right) \pi_2^B \tag{11}$$

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 her observed productivity in the first period.

Similarly promoting a minority worker is the best reply if (11) holds with reversed inequality sign. If expected productivity 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. Given that the employer of type low has no tastes for discrimination, it seems intuitive to assume that the employer breaks the tie in favor of the worker who displayed the highest productivity in the first period and if also productivity in the first period was the same, he tosses a coin. Using the population interpretation, tossing a coin means that half of the indifferent employers promote A and half B.

Employers of type $m^F = d$ are characterized by tastes for discrimination. It follows straightforwardly from assumption 4 that regardless of any observed and expected productivity level of the two workers:

$$BR^{d,F}\left(\mu^{d,F}\right) = \{\alpha = 1\}$$

a2) Workers.

Each worker, knowing her own type chooses the strategy s that maximizes her expected utility given her conjectures about opponents type strategy profile. Optimal actions in the second period conditioned on type and promotion have already been shown in assumption 3.

Substituting in the utility functions (1) and (2) the optimal values of effort in the second period and the possible value of the parameter m, we obtain:

$$U^{m^{\mathsf{A}}=1} = e_1^{m^{\mathsf{A}}=1} - \left(e_1^{m^{\mathsf{A}}=1}\right)^2 + \frac{\alpha K}{4} + \frac{(1-\alpha)}{4} \tag{12}$$

$$U^{m^{\mathsf{B}}=1} = e_1^{m^{\mathsf{B}}=1} - \left(e_1^{m^{\mathsf{B}}=1}\right)^2 + \frac{\alpha}{4} + \frac{K}{4}(1-\alpha) \tag{13}$$

$$U^{m^{\mathsf{A}}=K} = e_1^{m^{\mathsf{A}}=K} - \left(e_1^{m^{\mathsf{A}}=K}\right)^2 + \frac{\alpha}{4} + \frac{(1-\alpha)}{4K} \tag{14}$$

$$U^{m^{\mathsf{B}}=K} = e_1^{m^{\mathsf{B}}=K} - \left(e_1^{m^{\mathsf{B}}=K}\right)^2 + \frac{\alpha}{4K} + \frac{(1-\alpha)}{4} \tag{15}$$

As far as first period is concerned, l can be shown to be a strictly dominated action for all workers. The utility of a type $m^A = K$ of population A in the first period choosing i and h is respectively:

$$U^{m^{\mathsf{A}}=1}(i) = \frac{1}{4} + \left(\mu(\alpha=1|i)\frac{K}{4} + (1-\mu(\alpha=1|i))\frac{1}{4}\right)$$

$$U^{m^{\mathsf{A}}=1}(h) = \frac{K}{2} - \frac{K^2}{4} + \left(\mu(\alpha=1|h)\frac{K}{4} + (1-\mu(\alpha=1|h))\frac{1}{4}\right)$$

where $\mu(\alpha = 1|i)$ is the probability to be promoted that this worker thinks to have when playing i, and $\mu(\alpha = 1|h)$ is the probability to be promoted that this worker thinks to have when playing h.[18] Therefore, type $m^A = 1$ will choose h in the first period if

$$U^{m^{\mathsf{A}}=K}(e_1^{m^{\mathsf{A}}=K} = h) - U^{m^{\mathsf{A}}=K}(e_1^{m^{\mathsf{A}}=K} = i) > 0$$

which can be shown to be equivalent to

$$\mu(\alpha = 1|h) - \mu(\alpha = 1|i) > (K-1)\,. \tag{16}$$

Similarly,

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

It deserves to be mentioned that when there is no incentive to be promoted, the left hand side of these equation vanishes and therefore, given $K > 1$, h cannot be an optimal choice.

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

3.1.1 BNE: beliefs are correct

For all the players $\tilde{n}$ of every type m of each population the subjective probability distribution over opponents' type-strategy set coincides with the objective one. For instance, for the type m of population A :

$$\mu^{m,A}(m^B, s^B, m^F, s^F) = pr(m^B, s^B, m^F, s^F).$$

Intuitively, this means the probability that each player assigns to every combination of opponents' type-strategy profile is correct.

3.1.2 SCE: beliefs are not contradicted by the evidence

The distribution of observable outcomes $\hat{\sigma} = (\hat{\pi}_1^A, \hat{\pi}_1^B, \hat{\alpha}, \hat{\pi}_2^A, \hat{\pi}_2^B)$ obtained from all the stage games in a given round, coincides with the distribution that all the players $\tilde{n}$ of every type m of each population expect.

In other words, the objective probability to observe each of the finite number of elements σ must be equal to the subjective probability assigned by each player given her beliefs on the whole type strategy set. More intuitively, the subjective probability to observe a given element $\sigma = \left(\pi_1^A, \pi_1^B, \alpha, \pi_2^A, \pi_2^B\right)$ is obtained summing up the probability of all the type-strategy profiles that are mapped into terminal histories equal to σ. Formally,

$$\begin{aligned}
\Pr\{\sigma\} &= \sum_{m,s:\to\sigma} \mu_{\tilde{n}}^{m,A}(m,s) && (17)\\
&= \sum_{m,s:\to\sigma} \mu_{\tilde{n}}^{m,B}(m,s) && (18)\\
&= \sum_{m,s:\to\sigma} \mu_{\tilde{n}}^{m,F}(m,s) && (19)
\end{aligned}$$

The assumption of a continuum of players within each cohort make it unnecessary to consider explicitly in the subjective probability what each player knows, i.e. her own type and the chosen strategy, because the single observation is negligible.

3.2 Existence of the equilibria

This section focuses on the role of workers expectations while all the other potential causes of unequal outcome are neutralized. The distribution of types within the two populations of workers has already been assumed to be identical (assumption 5). Moreover:

Assumption 8: *beliefs of employers are correct* $\mu^F(\cdot) = pr(\cdot)$.

Assumption 9: *beliefs of workers A are correct* $\mu^A(\cdot) = pr(\cdot)$.

Assumption 10: *beliefs of workers B concerning the type-strategy profile of other workers B and of workers A are correct* $\mu^B(m^A, s^A) = pr(m^A, s^A)$.

Assumption 11: *beliefs of workers B about employers' strategies are correct* $\mu^B(s^F|m^F) = pr(s^F|m^F)$. In other words, workers B guess correctly the equilibrium strategy of each type of employer.

Assumption 12: *all workers B share the same beliefs about employers' type.*

Hence, the only source of heterogeneity between workers A and workers B is given by their expectations about employers' type. In particular, beliefs of workers B about employers' type may be correct $\mu^B(m^F) = pr(m^F)$ or wrong $\mu^B(m^F) \neq pr(m^F)$. Proposition 1 and Proposition 2 contrast what happens in these two different situations, everything else equal. $pr(m^F = 0)$ is the percentage of non discriminatory employers and $pr(m^F = d) = 1 - pr(m^F = 0)$ the percentage of discriminatory employers.

Proposition 1 *When workers B expectations about employers type are correct, i.e. $\mu^B(m^F) = pr(m^F)$, a 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 $1 - 0.5pr(m^F = 0)$.

In a BNE all beliefs are correct, therefore $\mu(\alpha|\cdot)$ are substituted with $pr(\alpha|\cdot)$. These probabilities must be the same within populations because workers cannot be distinguished, while they can differ across populations (for example because there are discriminatory employers). Conditions for the convenience of workers to supply h can be rewritten:

$$\left[pr(\alpha = 1|e_1^A = h) - pr(\alpha = 1|e_1^A = i)\right] > (K-1) \quad \text{for } m^A = 1 \quad (20)$$

$$\frac{1}{K}\left[pr(\alpha = 1|e_1^A = h) - pr(\alpha = 1|e_1^A = i)\right] > (K-1) \quad \text{for } m^A = K \quad (21)$$

$$\left[pr(\alpha = 0|e_1^B = h) - pr(\alpha = 0|e_1^B = h)\right] > (K-1) \quad \text{for } m^B = 1 \quad (22)$$

$$\frac{1}{K}\left[pr(\alpha = 0|e_1^B = h) - pr(\alpha = 0|e_1^B = h)\right] > (K-1) \quad \text{for } m^B = K \quad (23)$$

Since

$$\frac{1}{K}\left(pr(\alpha|h) - pr(\alpha|i)\right) < pr(\alpha|h) - pr(\alpha|i) \quad (24)$$

there cannot be an equilibrium in which worker $e_1^{m=K} = h$ and $e_1^{m=1} = i$.

Maximization of the non discriminatory employer

A good candidate for the equilibrium as far as employer strategies are concerned is:[19]

$$\begin{aligned} i, i &\rightarrow 0,5 \\ i, h &\rightarrow 0 \\ h, i &\rightarrow 1 \\ h, h &\rightarrow 0,5 \end{aligned} \quad (25)$$

where the action is defined as "percentage of workers A promoted" and for example h, i stands for worker A shows high productivity and worker B shows

[19] Couple of productivity levels where effort l is involved are not considered because l is strictly dominated for both workers.

intermediate productivity. We are considering a situation in which both types of population A choose the same action of the corresponding type of population B. There are 3 possible situations: a) all the workers choose h, b) all the workers choose i, c) $e_1^{m=K} = i$ and $e_1^{m=1} = h$.

Given that the distribution of types within populations is the same by assumption 5 and that beliefs about the distribution of opponents' type are correct, in a) and b) the employer would be certainly indifferent. In c) if the employer faces i, i or h, h he is indifferent, while if he faces h, i or i, h it is convenient to promote the worker of type H who supplied the higher productivity.

In all the situations a), b) and c), if there are not discriminatory employers, the candidate equilibrium strategy for non discriminatory employers implies that

$$\begin{aligned} pr(\alpha &= 1|e_1^A = h) - pr(\alpha = 1|e_1^A = i) = 0.5 \\ pr(\alpha &= 0|e_1^B = h) - pr(\alpha = 0|e_1^B = i) = 0.5 \end{aligned}$$

If not all the employers are non-discriminatory, i.e. if $pr(m^F = 0) < 1$

$$pr(\alpha = 1|e_1^A = h) - pr(\alpha = 1|e_1^A = i) = 0.5pr(m^F = 0) \tag{26}$$

$$pr(\alpha = 0|e_1^B = h) - pr(\alpha = 0|e_1^B = i) = 0.5pr(m^F = 0) \tag{27}$$

meaning that incentives to supply h are the same for both populations regardless of the percentage of discriminatory employers. This is intuitive, because the assumption that discriminatory employers always promote A implies that in both populations workers weight the convenience to supply h according to the percentage of non discriminatory employers. In the limit situation where there are only discriminatory employers, in both population promotion stops to be an incentive device, because A are sure to be promoted, B are sure not to be promoted.

Condition for the convenience to supply h, i.e. equations (20)-(23) above, become:

$$\begin{aligned} 0.5pr(m^F &= 0) > (K-1) \quad \text{workers } m^A, m^B = 1 \\ \frac{1}{K}0.5pr(m^F &= 0) > (K-1) \quad \text{worker } m^A, m^B = K. \end{aligned}$$

Obviously, 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 situations emerge, all representing a BNE with the characteristics 1) and 2) consistent with the candidate equilibrium strategy for the employer proposed in (25).

- If $K \geq 1 + 0.5pr(m^F = 0)$ all workers supply i. Non discriminatory employers promote 0.5 of A and 0.5 of B.
- If $\frac{1+\sqrt{1+2pr(m^F=0)}}{2} < K < 1 + 0.5pr(m^F = 0)$ workers $m = K$ of both populations supply i and workers $m = 1$ of both populations supply h. Non discriminatory employers promote 0.5 of A and 0.5 of B in i, i and h, h;they promote only A in h, i and only B in i, h.

- If $1 < K \leq \frac{1+\sqrt{1+2pr(m^F=0)}}{2}$ all workers supply h. Non discriminatory employers promote 0.5 of A and 0.5 of B.

For every $K > 1$ the fraction of workers A who are promoted is $1-0.5pr(m^F = 0)$. In the second period and in both populations workers $m = 1$ who are promoted supply h, workers $m = K$ who are promoted and workers $m = 1$ who are not promoted supply i, workers $m = K$ who are not promoted supply l.

Proposition 2 *There are $K > 1$ such that a Self-Confirming Equilibrium exists where*

1) all workers A supply h and all workers B supply i

2) only workers A are promoted.

due to wrong beliefs of minority workers which overestimate the percentage of discriminatory employers.

Conditions for the convenience of workers A to supply h do not change with respect to (20) and (21). As far as workers B are concerned:[20]

$$\left[\mu^B(\alpha = 0|e_1^B = h) - \mu^B(\alpha = 0|e_1^B = h)\right] > (K-1) \quad \text{for } m^B = 1 \quad (28)$$

$$\frac{1}{K}\left[\mu^B(\alpha = 0|e_1^B = h) - \mu^B(\alpha = 0|e_1^B = h)\right] > (K-1) \quad \text{for } m^B = K \quad (29)$$

The same candidate for the equilibrium as far as employer strategies are concerned is used:

$$\begin{aligned} i, i &\rightarrow 0,5 \\ i, h &\rightarrow 0 \\ h, i &\rightarrow 1 \\ h, h &\rightarrow 0,5 \end{aligned} \quad (30)$$

We are considering a situation in which one of the following holds all the workers of one populations supply i and all the workers of the other population supply h. Given that the distribution of types within populations is the same by assumption 5 and that beliefs about the distribution of opponents' type are correct, the employer is indifferent about the worker to promote and breaks the tie in favor of the worker who supplied the higher productivity in the first period.

The candidate equilibrium strategy for non discriminatory employers together with assumption 4 imply that:

$$pr(\alpha = 1|e_1^A = h) - pr(\alpha = 1|e_1^A = i) = 0.5pr(m^F = 0) \quad (31)$$

$$\mu^B(\alpha = 0|e_1^B = h) - \mu^B(\alpha = 0|e_1^B = i) = 0.5\mu^B(m^F = 0) \quad (32)$$

[20] Also in this case there cannot be an equilibrium in which worker $e_1^{m=K} = h$ and $e_1^{m=1} = i$. Given that beliefs are common within populations (assumption 12) it holds that

$$\frac{1}{K}\left[\mu^B(\alpha|h) - \mu^B(\alpha|i)\right] < \mu^B(\alpha|h) - \mu^B(\alpha|i).$$

Combining (21), (22), (28), (29), (31) and (32):

$$0.5pr(m^F = 0) > (K-1) \quad \text{for } m^A = 1 \tag{33}$$

$$\frac{1}{K}0.5pr(m^F = 0) > (K-1) \quad \text{for } m^A = K \tag{34}$$

$$0.5\mu^B(m^F = 0) > (K-1) \quad \text{for } m^B = 1 \tag{35}$$

$$\frac{1}{K}0.5\mu^B(m^F = 0) > (K-1) \quad \text{for } m^B = K \tag{36}$$

According to the parameter K different situations emerge.

If $K \geq 1 + 0.5\mu^B(m^F = 0)$ all workers B supply i.

If $1 < K \leq \frac{1+\sqrt{1+2\mu^B(m^F=0)}}{2}$ all workers B supply h.

If $K \geq 1 + 0.5pr(m^F = 0)$ all workers A supply i.

If $1 < K \leq \frac{1+\sqrt{1+2pr(m^F=0)}}{2}$ all workers A supply h.

When

$$1 + 0.5\mu^B(m^F = 0) \leq K \leq \frac{1 + \sqrt{1 + 2pr(m^F = 0)}}{2} \tag{37}$$

players' utility maximization implies the characteristics 1) and 2) of the proposition. All A supply h, all B supplying i;only workers A are promoted.[21] A necessary condition for this inequalities to hold is exactly $\mu^B(m^F = 0) < pr(m^F = 0)$, i.e. that workers B overestimate the percentage of discriminatory employer.

Empiricism

Workers A and employers have correct beliefs about other players' type-strategy profiles. Hence, the objective distribution of observables coincide with their subjective distribution. Workers B have wrong beliefs about employers' type only. Assumption 10 and 11 implies that their expected distributions of productivity (and therefore wages) within populations in the first period are correct. Associated to the observed outcome "workers A supply h - workers B supply i" their correct beliefs about employers strategies are associated to no worker B promoted. Finally, the expected distribution of wages within population in the second period is also correct.

Assumptions 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

[21] For istance, the inequalities are satisfied when:

- $\mu^B(m^F = 0) \leq 0.2$ and $pr(m^F = 0) \geq 0.25$ when $K = 1.1$
- $\mu^B(m^F = 0) \leq 0.4$ and $pr(m^F = 0) \geq 0.5$ when $K = 1.2$
- $\mu^B(m^F = 0) \leq 0.6$ and $pr(m^F = 0) \geq 0.8$ when $K = 1.3$.

to be discriminated against leading to unequal outcomes with only workers A promoted.[22]

Proposition 1 and 2 do not claim about uniqueness of BNE and SCE, respectively. Other BNE equilibria are for example associated to strategies of the employers different from (25). Outcomes of these equilibria can differ from those characterized above. However, all the BNE are symmetric, meaning that for every equilibrium with more than $1 - 0.5pr(m^F = 0)$ workers A promoted, there exists also another equilibrium in which less than $1 - 0.5pr(m^F = 0)$ workers A are promoted *ceteris paribus*.[23]

4 Conclusion

The aim of this paper was to set up a model where preferences and beliefs of both sides of the labor market matter. A framework is obtained where the main contributions to the discrimination literature can be nested. Moreover, the role of workers' expectations, neglected in the literature, can be analyzed.

Workers' expectations turn out to be very important, contributing to explain the puzzling long run persistence of cumulative discrimination. Even in a labor market where discriminatory tastes have disappeared, statistical discrimination is absent and all the other sources of heterogeneity such as distribution of ability among workers, etc. have been neutralized, unequal outcomes may arise due to workers' expectations. In this circumstance what happens is that in equilibrium wrong beliefs which overestimate the percentage of discriminatory employers are self-confirming. The reason is that minority groups who are (or expect to be) discriminated against supply less effort on average, because of a lower expected return. In other words, worker expectations to be discriminated against reduce the effectiveness of promotion as an incentive device. This induces a lower percentage of promotions within minority workers, which in turn convinces them that there are employers with discriminatory tastes. Moreover, in such a situation trial work periods, which can be an effective policy tool to break down statistical discrimination outcomes, are not an effective if workers have expectations of employers' discriminatory tastes.

[22] Observing only workers A promoted is certainly a knife-edge result. It is due to the strict assumptions made throughout the paper. Having more than two types of workers, for instance, makes it possible to observe SCE in which the fraction of workers A promoted is greater than that of the BNE but lower than 1.

[23] Characteristics of equilibria different from those proposed in Proposition 1 has been analyzed by means of simulations. Grid search do not display other SCE besides those described in Proposition 2 when $pr(m^F = 0) - \mu^B(m^F = 0) \geq 0.1$ under assumptions 1-12. Other SCE arise when $\mu^B(m^F = 0)$ is lower but close to $pr(m^F = 0)$. Outcomes of these equilibria do not differ significantly from those of Proposition 1 or Proposition 2. Many other SCE arise instead relaxing assumption 11, i.e. when beliefs of workers B about employer's strategies are not restricted to be correct. These SCE may be symmetric or not.

References

Aigner D.J. and G.G. Cain (1977), "Statistical Theories of Discrimination in Labor Markets", Industrial and Labor Relations Review, 30(2), January, 175-87.

Arrow K.J. (1973), "The theory of discrimination" in O. Ashenfelter and A. Rees (eds) "Discrimination in labor markets" pp. 3-33; Princeton University Press.

Battigalli P. and D. Guaitoli (1997) "Conjectural Equilibria and Rationalizability in a Game with Incomplete Information" in P.Battigalli, A.Montesano and F.Panunzi (eds) "Decisions, Games and Markets" pp. 97-124; Dordecht: Kluwer Academic Publishers.

Becker G.S. (1957), "The economics of discrimination"; University of Chicago Press.

Becker G.S. (1985), "Human Capital, Effort and the Sexual Division of Labor", Journal of Labor Economics 3, suppl. (January 1985): S33-S58.

Blau F.D. and C.L. Jusenius (1976), "Economists' Approaches to Sex Segregation in the Labor Market: An Appraisal." In Blaxall and Reagan (eds) "Women and the Workplace"; University of Chicago Press pp. 181-99.

Blau F.D. (1984), "Discrimination against women: theory and evidence" in W. Darity jr. (ed) "Labor economics: modern views" pp.53-89; Kluwer Nijhoff Publising.

Blau F.D. and M.A. Ferber (1991), "Career Plans and Expectations of Young Woman and Men", The Journal of Human resources, Vol 6 N.3, pp.277-301.

Breen R. and C. Garcia-Penalosa (1998), "Bayesian learning and Gender Segregation", mimeo.

Cain G.G. (1986), "The Economic Analysis of labor Market Discrimination: A Survey" in O. Ashenfelter and R. Layard (eds) "Handbook of labor economics" pp. 693-785; North Holland.

Coate S. and G.C. Loury (1993), "Will Affirmative-Action Policies Eliminate Negative Stereotypes?" American Economic Review Vol 83 N.5, pp. 1220-40.

Ferber M.A. and H.M. Lowry (1976), "The Sex Differential in Earnings: A Reapprisal", Industrial and Labor Relations Review, 29, April, 377-87.

Fudenberg D. and D.K. Levine (1998), "The Theory of Learning in Games"; Cambridge and London: MIT Press.

Mincer J. and S. Polacheck (1974), "Family investments in Human Capital: Earnings of Women", Journal of Political Economy, 82(2), 76-108.

Myrdal G. (1944), "An American dilemma", Harpher & Brothers.

Piore M.J. (1970), "Jobs and Training", in S.H. Beer (ed) "The State and the Poor" pp. 53-83; Cambridge, Mass.: Whintrop Press.

Phelps (1972) E.S. (1972), "The Statistical Theory of Racism and Sexism", American Economic Review, Vol 62 N.4, pp. 659-661.

Spence M.A. (1973), "Job Market Signaling", Quarterly Journal of Economics, Vol 87 N.3, pp. 355-374.

Spence M.A. (1974), "Market Signaling"; Cambridge, Mass.: Harvard University Press.