EUROPEAN UNIVERSITY INSTITUTE
Department of Economics

Discrimination and Workers' Expectations

Antonio Filippin

Thesis submitted for assessment with a view to obtaining the degree of Doctor of the European University Institute

Florence
December 2003

B/C →

EUROPEAN UNIVERSITY INSTITUTE
Department of Economics

Discrimination and Workers' Expectations

Antonio Filippin

The Thesis Committee consists of:

Pierpaolo Battigalli, Supervisor, Università Bocconi
Andrea Ichino, Internal Advisor, EUI
Peter Norman, University of Wisconsin
Yves Zenou, University of Southampton

Discrimination and Workers' Expectations

Antonio Filippin

01/12/2003

Contents

[1]This chapter is part of a joint project with Andrea Ichino.

Acknowledgements

In the process of developing this dissertation, I owe above all great gratitude to Pierpaolo Battigalli and Andrea Ichino for their excellent guidance. I think I was particularly lucky to have them as supervisors, not only for their indisputable skills, but also because their skills are a perfect mix. From their enthusiastic support in finding and developing the idea that became the core of the dissertation, to the stimulus of patiently extending the wall, brick after brick, over lands I did not know, their invaluable help made it possible for me to write a dissertation that now ranges from the theory to the evidence, both in the laboratory and in the field.

Besides the excellent feedback of the other members of the thesis committee, Peter Norman and Yves Zenou, this dissertation also benefited in different stages of its development from the valuable suggestions of Thomas Bauer, Giuseppe Bertola, Gary Charness, Roger Farmer, Ramon Marimon, Massimo Motta, Giuliana Palumbo, Karim Sadrieh and Karl Schlag. Comments from colleagues at the labor lunch at the EUI, from seminar participants at the University of Milan, University of Modena, University of Siena, and the European Commission - Employment and Social Affairs, as well as conference participants at the Economic Science Association 2003, European Association of Labour Economists 2003, Royal Economic Society 2003, European Economic Association 2002, IZA Summer School in Labor Economics 2002 and the Italian Association of Labor Economists 2002 and 2003 are also gratefully acknowledged. Needless to say, all the remaining errors are mine.

I would like to thank Daniele Checchi for the support he gave me in running the experiment in Chapter 3 and Bocconi University for allowing us to access the data used in Chapter 4. Giorgio Brunello, Claudio Lucifora and Rudolf Winter-Ebmer are gratefully acknowledged for allowing us to access some descriptive statistics of the data they collected. I would also like to thank Jacqueline Bourgonje, Marcia Gastaldo and Jessica Spataro for their assistance during the whole Ph.D. program and Nicola Hargreaves for her great job correcting my English.

Although not always easy, there is no doubt that this period was very important not only in terms of learning and work but also in terms of life experience. Besides giving me the opportunity to become once again a passionate soccer supporter and a (very poor, indeed) tennis player, Florence also allowed me to meet lots of new people, some of whom became very good friends of mine, like my neighbors at via Mimmole and my classmates, in particular Nicola Giammarioli, Thomas Steinberger and Daniela Vuri. A special thank you goes to Elena Argentesi, Massimo Motta and Alessandra Pierucci, who are probably unaware of how important their discussions about many unknown writers were in reminding me both of the importance of reading things without JEL codes on the front page, and that searching for equilibria is no less important in real life than in game theory.

Florence is also the place where my son Giacomo was born. My warmest and most special thanks goes to him, not only because he has already learned to make better use of a book that I bought during my first year of the Ph.D. program, but also, and above all, for the joy he gives me every single time I look at him. Finally, I want to thank Francesca, because, although at a crossroads we decided to proceed in different directions, I have not forgotten the long road we walked together.

Part I

Preface

General Summary

The neoclassical theory of discrimination is almost entirely a demand-side theory. There are very few contributions where workers' heterogeneity matters, and even fewer that study the possibility that unequal outcomes may arise or persist for reasons attributable to workers' expectations. This dissertation is aimed at filling this gap, analyzing the role that workers' expectations may play in explaining the long-run persistence of unequal outcomes that characterize some minorities in the labor market.

The first chapter introduces a generalization of the Self-Confirming Equilibrium concept to the case of aggregate observable outcomes, i.e. when players at the end of each round observe also the outcomes of games in which they are not directly involved. Such an extension becomes necessary to characterize the equilibria of a game, like that presented in Chapter 2, where social learning is assumed to be more important than individual outcomes directly observed by the agents. Moreover, aggregate outcomes ensure that Self-Confirming Equilibria do not rely on agents accessing different information.

The second chapter presents a model that analyzes the role of minority workers' expectations of being discriminated against. The model is formalized as a two-stage game of incomplete information in which populations of workers and employers are engaged, and where the preferences and beliefs of both sides of the labor market matter. 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 effort, which is a function of unobservable taste for work. Crucially, promotions depend via effort on workers' expectations about the unknown employer's type, 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 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 are unbiased and 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. What happens is that wrong beliefs of being discriminated against are self-confirming in equilibrium. Minority groups who expect 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.

The third chapter tests in the laboratory the predictions of the model in Chapter 2. The experiment replicates the model using a game where participants are randomly divided into two populations: Reds and Blues. In every trial each participant has an endowment of 10 euro cents and can decide how many cents to bid to get a prize worth 25 euro cents. Bets are not given back to the players, neither to the winner nor to the loser, making the game equivalent to an all-pay auction. The prize is awarded to the higher bid and it is split if bids are equal, unless the opponents are assigned to a "crazy computer" which instead awards the prize to the red player regardless of the bids. The mechanism underlying the Self-Confirming equilibrium driven by wrong beliefs is tested repeating the same treatment, where the fraction of crazy computers was equal to zero but unknown to the players, both before and after they face an increasing and known fraction of crazy computers. A reduction of bids made by disadvantaged subjects leading to a lower fraction of prizes assigned to them is observed in three out of seven sessions, in line with the theoretical predictions, but it vanishes rather quickly during the treatment, failing to generate the Self-Confirming Equilibrium driven by wrong beliefs. Stronger evidence supporting this Self-Confirming Equilibrium emerges from the strategy method, where, reacting to *ad hoc* aggregate statistics showing a decreasing fraction of prizes won by them, the blue players bid less and less. In turn, their lower bids make them less likely to win, confirming in some sense their wrong expectations that they were less likely to get the prize.

The fourth chapter tests on the field whether the expectations about the gender wage gap are in line with the realizations. Data collected among second year students of Bocconi University convey information about their wage expectations. Detailed controls allow a clean matching with a sample of Bocconi graduates providing information about their actual wages. The evidence shows that the gender gap implied by students' expectations one year after graduation is consistent with the gender gap implied by the earnings of their elder counterparts. There is instead a misperception of the gender gap later in the career because students expect the gender gap to be roughly constant while realizations indicate an increasing gap with experience, particularly for the relatively less skilled workers. Interestingly, there is also evidence that the realized gender gap at the beginning of a career is particularly high in the most recent cohorts and lower in the previous ones. Finally, our results suggest that the careers of females are characterized by "glass ceilings," in particular at high skill levels, and by "sticky floors" at the opposite end of the skill spectrum.

Results suggest that from a theoretical point of view workers' expectations can play an important role in explaining the long-run persistence of observed unequal outcomes that characterize some minorities in the labor market. An experiment shows that, to some extent, this is also the case in the laboratory, while field data provide evidence about how clearly the gender wage gap is perceived by a sample of students.

Extended Summaries

Chapter 1

Summary

This chapter provides an extension of the Self-Confirming (or Conjectural) Equilibrium concept to the case of aggregate observable outcomes within an extensive form game of incomplete information. The motivation for such an extension is that, in some circumstances such as the theoretical framework presented in Chapter 2, social learning is more informative than individual outcomes directly observed by the agents. Hence, player would choose to use aggregate observables, if available, instead of individual observables to form their beliefs. This happens in at least two situations: First, whenever players access at a larger scale the same relevant information available at the individual level; second, when individual observables are more informative, but each player participates to a sufficiently small number of repetitions of the game.

Chapter 2

Abstract

This chapter 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 groups of workers with the same characteristics 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.

Summary

This chapter explores the role played by minority workers' expectations of being discriminated against as an original explanation for the long-run persistence of observed unequal outcomes in the labor market. The main result is that in equilibrium groups of workers that have the same innate characteristics may be characterized by unequal outcomes. The reason is that minority workers' wrong expectations of being discriminated against become self-confirming. The idea is that minority groups who expect 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.

The model is formalized as a 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. Crucially for the results of this chapter, promotions depend via effort on workers' expectations about the unknown employer's type, which captures his possible disutility of promoting a minority worker.

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 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, as when information cascades occur, his observation would have a negligible information value. Moreover, the mechanism behind such a self-confirming equilibrium is robust both to trial work periods, which are instead an effective policy device to break down statistical discrimination outcomes, and to affirmative actions like quotas. The conclusion is that, from the theoretical point of view, workers' expectations may well contribute to explain the puzzling long-run persistency of unequal outcomes observed in the labor market.

Chapter 3

Abstract

This chapter is an experimental analysis of the role played by workers' expectations in explaining the puzzling long-run persistence of observed discrimination against certain minorities in the labor market. The experiment provides some evidence supporting the theoretical prediction that unequal outcomes may emerge due to disadvantaged workers' wrong expectations of being discriminated against. However, this effect is not long-lasting, since players learn the true state of nature in later stages of the experiment, failing to generate a Self-Confirming Equilibrium driven by wrong beliefs. The strategy method provides additional evidence that expectations matter.

Summary

This chapter is aimed at testing the predictions of the model in Chapter 2 and in particular that workers' expectations of being discriminated against may be an original explanation for the puzzling long-run persistence of observed discrimination against some minorities in the labor market.

The experiment replicates the model using a game where participants are randomly divided into two populations: Reds and Blues. In every trial each participant has an endowment of 10 Euro cents and can decide how many cents to bid to get a prize worth 25 Euro cents. Bets are not given back to the players, neither to the winner nor to the loser, making the game equivalent to an all-pay auction. Therefore, at the end of the trial the winner gets 25 Euro cents plus the amount not bet, while the loser gets only the amount not bet. In every lottery there are only two participants, one from the red population and one from the blue population. The players know that they face only one opponent in every trial and that it is possible to face the same opponent more than once during the same treatment, but of course they do not know when, given that random assignment takes place at the beginning of each trial. The prize is awarded to the higher bid and it is split if bids are equal, unless the opponents are assigned to a "crazy computer" which instead awards the prize to the red player regardless of the bids.

The mechanism underlying the Self-Confirming equilibrium driven by wrong beliefs in the theoretical model is tested by means of an identical treatment, where the fraction of crazy computers is equal to zero but unknown to the players, proposed both before and after they face an increasing and known fraction of crazy computers. A reduction of bids made by disadvantaged subjects leading to a lower fraction of prizes assigned to them is observed in three out of seven sessions, in line with the theoretical predictions, but it vanishes rather quickly during the treatment, failing to generate the Self-Confirming Equilibrium driven by incorrect beliefs. The parallel of this finding in the labor market would be a situation in which minority workers, after having been discriminated against, expect that unfavorable conditions continue while biased employers have actually disappeared. Hence, minority workers reduce their effort, and accordingly they are promoted less frequently, but eventually they discover that biased employers have disappeared and balanced promotions across populations of workers are observed.

The main reason why the experiment does not provide strong evidence in favor of unequal outcomes driven by subjects' expectations, is that the experiment fails to separate the beliefs of the two populations. In the theoretical model different expectations about the fraction of discriminatory employers are a necessary condition for the Self-Confirming Equilibrium driven by wrong beliefs, but in the lab Blues never expect a clearly higher fraction of crazy computers than Reds. A possible explanation for this finding is that the experimental design overemphasizes the discontinuity between treatments, preventing carryover effects from emerging. Another possible explanation is that the set of choices is roughly continuous. This induces only a very slight difference in the optimal behavior of advantaged and disadvantaged subjects, and therefore also a very low cost of experimenting to discover the true state of nature. Finally, it is also worth noting that, from a regression where all observable and unobservable individual characteristics are controlled for, Blues display a propensity to bid much more than Reds *ceteris paribus* in spite of the random assignment of the color.

The strategy method, on the other hand, supports the Self-Confirming Equilibrium driven by wrong beliefs. In fact, advantaged and disadvantaged subjects react in a different way when *ad hoc* aggregate statistics are displayed. Subjects are asked to bid five times, after five different fictitious distributions of prize winners across populations in the previous period have been displayed. This fictitious distribution shows a fraction of prizes awarded to the Blues decreasing from 80% to 0% in subsequent trials. Although all subjects are informed that there is no computer that awards the prize according to the color label, what happens is that blue players are influenced by the aggregate statistics showing that a decreasing fraction of them gets the prize, and thus bid less and less. In turn, their lower bids make them less likely to win, leading to unequal outcomes, which are consistent with wrong expectations that they were less likely to get the prize.

The experiment also deals with the relevance of the information structure by dividing the two populations into two subgroups, one observing individual outcomes only (bids and winners of the games in which the player is directly involved), while the other is also informed about the distribution of bids within, as well as the distribution of promotion across, populations. From the experiment there is no evidence of different patterns between these two subgroups.

Concluding, findings of the experiment provide some evidence supporting the Self-Confirming Equilibrium driven by wrong beliefs, but the evidence cannot be considered fully satisfactory. In order to provide a more robust test of the theoretical model, future experiments should be modified to include such things as dichotomized alternatives or less clear-cut treatments.

Chapter 4

Abstract

This chapter explores the extent to which the gender wage gap is anticipated by workers' expectations. Data collected among second year students of Bocconi University convey information about their wage expectations. Detailed controls allow a clean matching with a sample of Bocconi graduates providing information about their actual wages. The evidence shows that the gender gap implied by students' expectations one year after graduation is consistent with the gender gap implied by the earnings of their elder counterparts. There is instead a misperception of the gender gap later in the career after graduation because students expect the gender gap to be roughly constant while realizations indicate an increasing gap with experience, particularly for the relatively less skilled workers. There is also evidence that the gender gap at the beginning of a career is particularly high in the most

recent cohorts and lower in the previous ones. Finally, our results suggest that the careers of females are characterized by "glass ceilings," in particular at high skill levels, and by "sticky floors" at the opposite end of the skill spectrum.

Summary

The evidence presented in this chapter points towards some interesting findings. We show that the gender gap implied by students' expectations one year after graduation is consistent with the gender gap observed in the actual earnings of their older counterparts who have already graduated. There is instead a misperception of the gender gap ten years after graduation because students expect the gender gap to be roughly constant while realizations point toward an increasing gap with tenure. The gender gap diminishes but does not disappear when several controls such as family background, place of birth, high school diploma, university program attended, performance at university, civil status and number of children are taken into account.

A second set of intriguing results concerns the evidence on realized gender gaps independently of expectations. Here, in contrast with the recent literature for industrialized countries, we see no evidence of a diminishing gender gap between subsequent cohorts of Bocconi graduates at the beginning of a career. In particular, the gender gap immediately after graduation displays a puzzling upward trend and reaches particularly high and significant values in the most recent 1997 cohort. This result is likely to be a consequence of the elimination of the cost of living adjustment scheme called *Scala Mobile* which prevailed in Italy during the '80s and was abolished in 1992.

Finally, while the gender wage gap for the best graduates is large already at the beginning of a career but remains more or less constant throughout the working life, for the worse graduates the gender gap starts slightly lower but increases more significantly with experience. These results suggest that the careers of females are characterized by “glass ceilings,” in particular at high skill levels, and by “sticky floors” at the opposite end of the skill spectrum. Unfortunately, our data do not allow us to shed more light on the real nature and on the determinants of these differences in career developments.

Part II

Discrimination and Workers' Expectations

Chapter 1

An Extension of the Self-Confirming Equilibrium Concept to the Case of Aggregate Observables

1.1 Introduction

The Conjectural (or Self-Confirming) Equilibrium captures a situation in which no player receives any signal which makes her change her beliefs about the type and strategy of the opponents. Therefore, players' beliefs can be wrong, as long as players do not receive information inconsistent with their wrong beliefs. In the literature the conjectural equilibrium concept is analyzed within a wide range of extensive form games of incomplete information. One example is given by infinitely repeated games with interactions among players who are always the same at every stage of the game and whose type is private information. Another example is given by the so-called population games, in which at every stage each player is randomly matched with one or more players of the other population(s) whose type is private information. In the latter example every player could face potentially different opponents at every stage. Therefore, her conjecture must concern the distribution of types and strategies within the population from which the potentially different opponents are chosen, and no longer the type-strategy of some fixed opponents. However, a common feature of these specifications is that players are assumed to observe (and keep track of) the outcomes of the games in which they are directly involved, and not of the aggregate outcomes, i.e. the outcomes of all the constituent games played in a given period within the given society.[1]

This chapter provides an extension of the Self-Confirming (or Conjectural) Equilibrium concept, as introduced by Battigalli (1987), Fudenberg and Levine (1993 and 1998) and Dekel Fudenberg and Levine (2003), to the case of aggregate observable outcomes within an extensive form game of incomplete information.[2] Battigalli and Siniscalchi (2003) also consider this case, but they assume that the whole joint distribution over the set of types and outcomes is observable, whereas I assume that only the distribution of terminal histories of the constituent games is observable.

The motivation for such an extension is that, in some circumstances, social learning is more informative than individual outcomes directly observed by the agents. Hence, players would choose to use aggregate observables, if available, instead of individual observables to form their beliefs. This happens in at least two situations: first, whenever players access at a larger scale the same relevant information available at the individual level; second, when individual observables are more informative, but each player participates in a sufficiently small number of repetitions of the game.

[1] "Constituent game," as defined for example in Osborne and Rubinstein (1994), is used instead of the more common "stage game" to identify every repetition of an infinitely repeated supergame. The reason is to avoid confusion between one period and one repetition of the game (which can be composed of many periods), since stage is used in both senses in the literature.

[2] In this chapter the generalization of the Self-Confirming Equilibrium concept is limited to to the particular information structure and the setting of the game characterizing the model presented in chapter 2.

In the first case, suppose that each player observes the history (i.e. the complete sequence of actions) of the game in which she is directly involved as well as all the statistical distribution of histories arising from all the other games where other players are involved. If private information does not convey more information than the history of the game, directly observed outcomes have a negligible information value with respect to the joint distribution of histories at the aggregate level. Therefore, players will form their beliefs using aggregate observables only.

In the second case, suppose that individual outcomes have an additional information value compared to the aggregate observables, for instance because at aggregate level only the marginal distributions of actions across populations are known, while players observe the history of the game in which they are directly involved. If agents play infinitely many times, individual outcomes will be more informative. However, if each agent participates in a limited number of repetitions of the game, aggregate outcomes may still be an (approximately) sufficient satistic to form her beliefs.

An example should be helpful to make the point. Chapter 2 describes a model that analyzes the role of workers' expectations of being discriminated against in order to explain unequal outcomes that characterize some minorities in the labor market. A game of incomplete information with observable actions and private values is presented, with players representing populations of workers and employers. In every constituent game, i.e. in every repetition of the game played by actors randomly drawn from their populations, three players participate: two workers, one of whom belongs to a minority group, and one employer. The employer promotes one (and only one) of the two workers after having observed their effort, which is a function of unobservable taste for work. Promotions depend via effort on workers' expectations about the unknown employer's type, which captures his possible disutility of promoting a minority worker. Players try to figure out the distribution of types and strategies among the populations of opponents using available information. In particular, minority workers try to figure out their probability of being promoted conditional on observed workers' output.

Within this framework, a history is a sequence of actions that includes the productivity of workers in the first period, the decision about promotion and the productivity of workers in the second period.

When the sequences of actions arising from all the constituent games are public domain, the information directly available from the games in which each player is directly involved becomes negligible because the probability of being promoted conditional on workers' output can easily be inferred from aggregate statistics (first case above).

Suppose instead that aggregate observables are not as informative as the individual observables, for example because only the marginal distributions of actions across populations are available. This means that the distributions of effort and promotions across populations are known, but in this case the exact combination of effort levels upon which each promotion has been decided cannot be retrieved. Instead, this information is available at the end of the constituent game in which each player is involved (the game is characterized by observable actions). Hence, the probability of being promoted conditional on pairs of effort levels is better inferred through individual observables, provided that a sufficiently large number of observations is available. If, instead, each player participates in a finite and small number of repetitions of the game, as is intuitively the case when considering the decision about her promotion that a worker might face during her career, aggregate observables might be relatively more informative (and this will certainly be the case at the beginning of her working life).

1.2 The structure of the game

What follows is a simple extensive form game of incomplete information which allows the Self-Confirming (or Conjectural) Equilibrium concept to be extended to the case of aggregate observable outcomes.

The model is formalized as an infinitely repeated extensive game of incomplete information in which N populations are engaged. The population game is repeated infinitely under random matching at the beginning of each period, meaning that at every repetition of the game every player is randomly matched with one player from each of the other populations to play the so-called **constituent game.** I use the following assumptions and notation.

- **Populations.** There are N populations characterized by a continuum of players. Each population $i = 1, ..., N$ corresponds to a role in the constituent game.

- **Incomplete information.** Every indvidivual player of each population knows her own payoff type only.

  The number of types is finite and the distribution of types within each population is assumed to be fixed. Hence, there is an infinite number of individuals for every role i and type θ_i, so that assuming that the law of large numbers holds, the actual fraction of any given combination of types coincides (almost surely) with its probability.

  $\Theta = \Theta_1 \times \Theta_2 \times ... \times \Theta_N$ is the set of type profiles;

  $\theta = (\theta_1, ..., \theta_N) \in \Theta$ is the vector of types characterizing the players in a constituent game;

  $p_i \in \Delta(\Theta_i)$ is the statistical distribution of types in population i.

  Types are statistically independent across populations. Thus $p = p_1 \times ... \times p_N \in \Delta(\Theta)$ is the true statistical distribution of type profiles, fixed once and for all at the beginning of the repeated interaction.

In every constituent game, the probability of facing a given combination of types of opponents $\theta_{-i} \in \Theta_{-i}$ coincides with the (fixed) product of the frequencies of each type of opponents within her population, that is $p_{-i}(\theta_{-i}) = \prod_{j \neq i} p_j(\theta_j)$.

- **Histories.** $z \in Z$ denotes a generic terminal history, i.e. a complete sequence of actions, played in a constituent game. $\sigma \in \Delta(Z)$ denotes a distribution of terminal histories in the population game, obtained pooling the terminal histories z arisen in all the constituent games. $h \in H$ denotes non-terminal histories.

- **Strategies.** Strategies in the constituent game are functions that map non-terminal histories into feasible actions $s_i : H \rightarrow A_i$, such that $s_i(h) \in A_i(h)$ for each $h \in H$, where $A_i(h)$ is the set of feasible actions for player/role i at history h. Note that players know their type from the very beginning of the game, i.e. there is no *ex ante* stage in which they plan how to play without knowing their type. This is the reason why strategies are not defined as functions of players' types.[3] $S = S_1 \times ... \times S_N$ denotes the set of strategy profiles. S_{-i} denotes the set of i's opponents' strategy profiles.

  Each player is assumed to play pure strategies. However, in the constituent game every player faces, in every role other than hers, a population of opponents playing pure strategies, which is equivalent to face one opponent playing a mixed strategy. The probability associated with each pure strategy in the support of such a mixed strategy coincides with the frequency of players playing that pure strategy.

[3] On the other hand, if there is an *ex ante* stage at which i plans how to play the game before she learns her type, it makes sense to define i's strategies as functions with domain $\Theta_i \times H$. Note that the *ex ante* and the *ex post* formulations of the Self-Confirming Equilibrium are not necessarily equivalent.

$\zeta : S \to Z$ is the function that in every constituent game maps players' strategies into terminal histories.

- **Payoffs.** $u_i : \Theta \times Z \to \mathbb{R}$ denotes agent i's payoffs, which depend both on the strategies played in the constituent game where she was directly involved and on her type. Note that the present formulation allows for interdependent values, since agent i's payoffs may also depend on opponents' type.[4]

Players try to figure out the distribution of types and strategies among the populations of opponents using available information. Besides observing the sequence of actions $z \in Z$ of the constituent game in which she is involved, every player is assumed to observe also $\sigma \in \Delta(Z)$, i.e. the distribution of sequences of actions which have occurred in all the constituent games played in a given period. In other words, each player observes not only the outcomes of the game in which she is directly involved but also the distribution of histories in the population game. Players are therefore assumed to access very informative statistics. Individual outcomes become a useless source of information when compared to available aggregate outcomes in order to infer the type-strategy profile of the opponents.

1.3 Definition

The Self-Confirming (or Conjectural) Equilibrium concept, applied to the framework described above, captures a situation in which:

a) each player maximizes her utility given her beliefs about the type-strategy profile of the opponents.

b) no player receives any signal that contradicts her beliefs.

For the sake of simplicity, it is assumed that individual players have point beliefs $\mu \in \Delta(\Theta \times S)$ about the actual distribution of types and strategies in

[4]When payoffs of an agent are not a function of the opponents' type, the game is characterized by private values.

the population game.[5] Assuming that players know that types are independent across populations, it makes sense to assume that in a steady state beliefs are represented by product measures: $\mu = \mu_1 \times ... \times \mu_N$, where $\mu_i \in \Delta(\Theta_i \times S_i)$ denotes the marginal of μ on $\Theta_i \times S_i$. The marginal of μ on $\Theta_{-i} \times S_{-i}$ is denoted $\mu_{-i} \in \Delta(\Theta_{-i} \times S_{-i})$. These beliefs refer to the general case of a game characterized by interdependent values. When a game is instead characterized by private values and no assumption is made concerning the knowledge of opponents' rationality, the Self-Confirming Equilibrium can be defined using beliefs about strategies only $\mu \in \Delta(S)$. Beliefs about types can be disregarded, unless players believe that opponents are rational. Knowledge of opponents' rationality make it necessary to define beliefs about types as well, because in this case players are aware that whether a strategy of an opponent is a best reply or not depends on her type. The model presented in Chapter 2 is characterized by private values, but it implicitly assumes that players believe that the opponents are rational. Hence, beliefs will be defined also about types.

Players form their beliefs about the type-strategy profile of the opponents using aggregate observables. However, both the role of the individual player in the constituent game (i.e. the population she belongs to) and her type may affect her beliefs. For instance, the dataset used in Chapter 4 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 its realization. However, they assign a different importance to the underlying causes. In fact, while a larger fraction of men thinks 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.

In the definition below it is assumed that all individual players of the same type θ_i within each population i, share the same beliefs, denoted $\mu^{\theta_i} \in \Delta(\Theta \times S)$, and

[5]Otherwise, beliefs would have the more general form $\mu \in \Delta[\Delta(\Theta \times S)]$.

play the same strategy, denoted $s_i^{\theta_i} \in S_i$. This assumption can be easily relaxed, at the cost of some notational complexity.

Definition 1 *A Self-Confirming Equilibrium is a profile of strategies and beliefs* $(s_i^{\theta_i}, \mu^{\theta_i})_{i=1,...,N,\theta_i\in\Theta_i}$ *such that for every population* $i = 1, ..., N$ *and every type* $\overline{\theta}_i \in \Theta_i$:

a) The chosen strategy $s_i^{\overline{\theta}_i}$ *is best reply to the marginal belief* $\mu_{-i}^{\overline{\theta}_i}$ *about the opponents' type-strategy profile:*

$$s_i^{\overline{\theta}_i} \in \arg\max_{s_i\in S_i} \sum_{(\theta_{-i},s_{-i})\in(\Theta_{-i}\times S_{-i})} \mu_{-i}^{\overline{\theta}_i}(\theta_{-i}, s_{-i})\, u_i(\theta_i, \theta_{-i}, \zeta(s_i, s_{-i})).$$

b) The observed distribution of outcomes σ*, obtained from all the constituent games, coincides with the subjectively expected distribution obtained from belief* $\mu^{\overline{\theta}_i}$*, that is, for every terminal history* $z \in Z$ *it must be the case that:*

$$\sum_{\theta:\zeta(s_1^{\theta_1},...,s_N^{\theta_N})=z} p(\theta) = \sigma(z) = \sum_{(\theta,s):\zeta(s)=z} \mu^{\overline{\theta}_i}(\theta, s).$$

It is worth noting that, while beliefs in condition a) refer to opponents only, beliefs in condition b) also include the population from which player i is drawn. The reason is that every player accesses aggregate observables, i.e. the distribution of terminal histories in the whole population game, where also other players from the same population as hers are involved.

Bibliography

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

[2] BATTIGALLI, P. AND M. SINISCALCHI (2003): "Rationalization and Incomplete Information," *Advances in Theoretical Economics,* forthcoming.

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

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

[5] ——— (1998): *The Theory of Learning in Games.* Cambridge and London: MIT Press.

[6] OSBORNE, M.J. AND A. RUBINSTEIN (1994): *A Course in Game Theory.* Cambridge: MIT Press.

Chapter 2

Discrimination and Workers' Expectations

2.1 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 chapter is to analyze the role of workers' expectations, so far neglected in the literature, in explaining the observed unequal outcomes that characterize some minorities in the labor market. The idea is that minority groups who expect 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.

The model is formalized as a 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 effort. Crucially for the results of this chapter, promotions depend via effort on workers' expectations about the unknown employer's type, which captures his possible disutility of promoting a minority worker.

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 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 to explain the puzzling long-run persistency of unequal outcomes observed in the labor market.

The structure of this chapter is as follows. After some definitions are outlined in Section 2.1.1, the constituent game of the model, i.e. the game after the players have already been matched, is presented in Section 2.2.1. The population game, the matching process and the information structure, necessary to characterize beliefs, are described in Section 2.2.2. Section 2.3 concentrates on the analysis of the equilibria of the model and its policy implications. The connections of the model to the related literature are sketched in Section 2.4. Section 2.5 concludes.

2.1.1 Definitions

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 workers that have the same charateristics *ex ante*, 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 im-

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

portant, 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 discrimination**, defined as different achievements for workers that have the same characteristics *ex ante*.

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.

2.2 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 effort (productivity) e. The observable characteristic distinguishes the so-called majority worker, identified by subscript A, from the so-called minority worker, identified by subscript B. Employers are denoted by subscript F. Workers compete in order to be promoted. Promotions depend on both employers' and workers' type as well as on their beliefs about the opponents' type-strategy profile. Crucially for the results of this chapter, promotions depend via effort on workers' expectations about the unknown employer's type, which captures his possible disutility of promoting a minority worker.

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.

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' effort 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 $e_A^2, e_B^2 \in E$ in the second period.

The constituent game is characterized by observable actions, because workers' effort as well as the decision about promotion are directly observed. The game is also characterized by incomplete information, because every player knows his or her type only.

Incomplete information

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

$\theta_A \in \Theta_A$ 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 (2.1) and (2.2) below). There are only two possible types of worker $\Theta_A = \Theta_B = \{1, K\}, K > 1$. The higher the parameter, the higher the cost of effort.[2]

[2]Taste for work is assumed to matter in the second period only, as if one of the two types of

$\theta_F \in \Theta_F$ represents employer's tastes for discrimination, i.e. the disutility of promoting a minority worker. For the sake of simplicity, the types of employer are restricted to $\Theta_F = \{0, d\}$, 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' effort.

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.[3]

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

$$U^{\theta_A} = w_A^1 - (e_A^1)^2 + w_A^2 - I(\alpha_A)\frac{\theta_A}{K}(e_A^2|\alpha_A)^2 - I(\alpha_B)\theta_A(e_A^2|\alpha_B)^2, \qquad (2.1)$$

$$U^{\theta_B} = w_B^1 - \left(e_B^1\right)^2 + w_B^2 - I(\alpha_A)\theta_B(e_B^2|\alpha_A)^2 - I(\alpha_B)\frac{\theta_B}{K}(e_B^2|\alpha_B)^2, \qquad (2.2)$$

where:

worker ($\theta = K$) became lazier after the promotion decision. Allowing different tastes for work in the first period would *de facto* resolve the 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 2.3. 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). It is worth noting that assuming that the type differs in the second period only does not imply that the first period is negligible. Although the istantaneous utility fuction is the same for all workers in the first period, the choice of effort depends on the utility during the whole working life. In particular, workers characterized by a lower cost of effort in the second period have a stronger incentive to exert a higher effort also in the first period, in order to increase their probability of being promoted.

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

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

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

α_A means that worker A is promoted, α_B that worker B is promoted;

$I(\cdot)$ is the indicator function that assigns the value 1 when the argument is true, for instance when the worker A is promoted $I(\alpha_A) = 1, I(\alpha_B) = 0$;

$\frac{1}{K} < 1$ 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. $w^t = e^t$. The labor market is assumed to be competitive, therefore productivity is entirely paid to workers.[4] 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 and employers compete on enforceable piece-rate contracts. Workers would be free to move, but in equilibrium $w^t = e^t$ and nobody moves. A similar argument can be used to justify inside promotions in equilibrium.[5]

It follows from the utility function in (2.1) and (2.2) that in the second period effort will be higher if the worker is promoted

$$\begin{aligned} e^{2*}_A|\alpha_A &= \frac{K}{2\theta_A} > e^{2*}_A|\,(\alpha_B) = \frac{1}{2\theta_A}, \\ e^{2*}_B|\alpha_B &= \frac{K}{2\theta_B} > e^{2*}_B|\,(\alpha_A) = \frac{1}{2\theta_B}. \end{aligned}$$

Substituting the type of workers $\theta_A, \theta_B \in \{1, K\}$ 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]

[4] In this model, productivity stands for output per worker. It does not refer to a worker's innate characteristic. Therefore, a more productive worker is simply a worker characterized by a higher output.

[5] This assumption 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 have been solved using numerical simulations, and it turns out to be much more complicated without being more insightful (see also footnote 2).

[6] The specification of the utility function adopted in the paper is the same as that proposed by

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." They coincide with the optimal choice in the second period of a good 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 bad type who is not 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

$$U^{\theta_F} = (m - 1)(e_A^1 + e_B^1 + e_A^2 + e_B^2) - \alpha_B \theta_F,$$

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 $\alpha_B \theta_F$ 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.

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 rather than an employer. Instead of profits, the supervisor maximizes a bonus which is a fraction of the overall productivity of the workers. Nothing would change in practice, because also the supervisor has the incentive to promote the more productive worker in order to maximize his\her bonus.

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 effort 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 her type, is simply $\{\alpha_A, \alpha_B\}$, where α_A stands for "promotes worker A" and α_B 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 effort levels. 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 different type, the employer maximizes her utility promoting the worker characterized by higher tastes for work (i.e. lower θ). Moreover, in order to prevent unequal outcomes from arising because of asymmetric choices of employers who are instead supposed to be indifferent to workers' membership, the following assumption is necessary.

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).

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.

2.2.2 The population game

The constituent game described in section 2.2.1 is inserted in a supergame, called population game, which specifies how players are matched and what information they can access. This supergame is not formally described in what follows. The reason is that the equilibria of the constituent game that will be presented in section 2.3 can be thought as rest points of a learning process, and therefore they are analyzed in a stationary framework. Since the main role of the population game is simply to provide a dynamic framework that justifies such equilibria, the description will focus only on the information structure that allows 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.

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 assumed 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.

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, $z = (e_A^1, e_B^1, \alpha, e_A^2, e_B^A) \in Z$ 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.

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 her own type and the strategy she chooses only, the unknown component of $\Theta \times S$ is 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 θ are defined as

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

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 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 update their beliefs observing workers' effort. Employers' prior beliefs are a probability measure over each worker's type-strategy profile $\mu^{\theta_F}(\theta_A, s_A), \mu^{\theta_F}(\theta_B, s_B)$. 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

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

the (strictly positive) probability that an employer of type θ assigns to the effort level $\tilde{e}_A^1 \in E$ according to his\her prior beliefs, updated beliefs after the observation of $\tilde{e}_A^1$ will be:

$$\mu^{\theta_F}(\theta_A, s_A | \tilde{e}_A^1) = \begin{cases} \frac{\mu^{\theta_F}(\theta_A, s_A)}{\mu^{\theta_F}(\tilde{e}_A^1)} & \text{if } \theta_A, s_A : e_A^1 = \tilde{e}_A^1 \\ 0 & \text{if } \theta_A, s_A : e_A^1 \neq \tilde{e}_A^1 \end{cases} .$$

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 in Chapter 4 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 its realization. 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:*

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

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 of being promoted can be signaled only through a higher effort. All the equilibria that could possibly arise when workers signal their willingness of being promoted supplying l are excluded. Moreover, exerting effort l in the first period becomes strictly dominated for all workers.

2.3 Analysis of the equilibria

In this section two qualitatively different sets of equilibria are presented. The first set (see Proposition 2) displays symmetric outcomes under the assumption that the expectations of all players are correct. The second set of equilibria (see Proposition 3) 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 Chapter 1 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 her beliefs, updated whenever possible, about every possible opponents' type-strategy profile (see section 2.3.1). 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 assumption, beliefs about opponents' types are correct as well. On the other hand, when the Common Prior assumption is not satisfied, beliefs may be contradicted by the evidence and the BNE could not be justified as a rest point of a learining process. 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 2 and Proposition 3 are Perfect Bayes-Nash (henceforth: PBNE) and Self-Confirming at the same time. This is fairly intuitive in Proposition 2 given that the Common Prior assumption is satisfied and therefore beliefs are correct. Nevertheless, this is the case also in Proposition 3 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 2.3.2).

2.3.1 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 her utility promoting the worker characterized by higher tastes for work (i.e. lower θ).

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

$$\alpha_A = BR^{0_F}\left(e^1\right) \text{ if}$$
$$\sum_{e_A^2 \in E} \mu^{0_F}(e_A^2|e_A^1)e_A^2 > \sum_{e_B^2 \in E} \mu^{0_F}\left(e_B^2|e_B^1\right) e_B^2, \tag{2.3}$$

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 2.3 holds with reversed inequality sign. If expected productivity in the second period is the same, the non-discriminatory employer is indifferent.[8] This means that both α_A and α_B as well as all the mixed strategies would be best replies. However, the only non-trivial strategy that satisfies 2.3, that implements a fair tournament (assumption 3) and that satisfies monotonicity (assumption 5) is:

$$i,i \rightarrow 0.5;\ h,i \rightarrow 1\ ;\ i,h \rightarrow 0;\ h,h \rightarrow 0.5, \tag{2.4}$$

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 in the first period are *high* and *intermediate*, respectively.[9]

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

[8]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).

[9]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.

All the employer's strategies that do not contain $h, i \rightarrow 1$; $i, h \rightarrow 0$ are weakly dominated.

any observed and expected productivity level of the two workers:

$$\alpha_A = BR^{d_F}(e^1), \forall e^1.$$

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 (2.1) and (2.2) we obtain:

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

As far as the first period is concerned, it can be shown that all the strategies containing l are stricly dominated for all workers as long as they believe that employers' strategies are monotone (assumption 5). The utility of a type $\theta_A = 1$ choosing i and h in the first period is, respectively:

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

where $\mu(\alpha_A|i)$ and $\mu(\alpha_A|h)$ are the subjective probabilities of being promoted when playing i and h, respectively.[10] Therefore, type $\theta_A = 1$ will choose h in the

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

first period if $U(h) - U(i) > 0$, which leads to

$$\mu(\alpha_A|h) - \mu(\alpha_A|i) > (K-1).$$

Similarly,

- a worker $\theta_B = 1$ will choose h if $\mu(\alpha_B|h) - \mu(\alpha_B|i) > (K-1)$;
- a worker $\theta_A = K$ will choose h if $\frac{1}{K}(\mu(\alpha_A|h) - \mu(\alpha_A|i)) > (K-1)$;
- a worker $\theta_B = K$ will choose h if $\frac{1}{K}(\mu(\alpha_B|h) - \mu(\alpha_B|i)) > (K-1)$.

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 $K > 1$.

2.3.2 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,

$$\mu^{\theta_A}(\theta_B, s_B, \theta_F, s_F) = \Pr(\theta_B, s_B, \theta_F, s_F) \tag{2.5}$$

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 the observed distribution of outcomes σ, obtained from all the constituent games, coincides with the subjectively expected distribution. The latter 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 (2.5) for some type θ or strategy s, are not contradicted by the evidence.[11]

2.3.3 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;

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 both of workers A and of 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 perturba-

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

tions 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 as well, or allowing more heterogeneity of fundamentals within the model.

Some of the assumptions made so far are very convenient from the theoretical point of view, because they allow 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.[12] It deserves to be stressed once more that such assumptions are made with the only purpose to focus 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 2 and Proposition 3 contrast what happens in these two different situations, everything else being equal.

Proposition 2 *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 $1 - 0.5\Pr(\theta_F = 0)$.

Sketch of proof

[12]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 2 and 3.

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:

$$\Pr(\alpha_A|h) - \Pr(\alpha_A|i) \geq K - 1 \qquad \text{for } \theta_A = 1 \tag{2.6}$$

$$\frac{1}{K}(\Pr(\alpha_A|h) - \Pr(\alpha_A|i)) \geq K - 1 \qquad \text{for } \theta_A = K \tag{2.7}$$

$$\Pr(\alpha_B|h) - \Pr(\alpha_B|i) \geq K - 1 \qquad \text{for } \theta_B = 1 \tag{2.8}$$

$$\frac{1}{K}(\Pr(\alpha_B|h) - \Pr(\alpha_B|i)) \geq K - 1 \qquad \text{for } \theta_B = K. \tag{2.9}$$

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

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

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 (2.4) as a candidate equilibrium strategy for the non-discriminatory employers. Equation (2.4) 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 2 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 (2.4) implies that for all workers

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

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

$$\Pr(\alpha|h) - \Pr(\alpha|i) = 0.5 \Pr(\theta_F = 0)\,. \tag{2.10}$$

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 (2.10) into equations (2.6)-(2.9) above, the conditions that make h the optimal choice can be rewritten

$$\begin{aligned} 0.5 \Pr(\theta_F = 0) &> K - 1 \text{ for workers } \theta_A, \theta_B = 1, \\ \frac{1}{K} 0.5 \Pr(\theta_F = 0) &> K - 1 \text{ for workers } \theta_A, \theta_B = K. \end{aligned}$$

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 2 and consistent with the candidate equilibrium strategy for the employer proposed in (2.4). 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 2 are not unique. For instance, there are other PBNE associated with strategies of the employers different from (2.4). 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 more than $0.5\,\Pr(\theta_F = 0)$ workers B are promoted.[13]

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 3 refers only to PBNE that are also SCE, i.e. when beliefs are not contradicted by the evidence.

Proposition 3 *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.

[13] Characteristics of equilibria different from those proposed in Proposition 2 have been analyzed by means of simulations.

Sketch of proof

For such an equilibrium to exist, conditions for the incentive of workers A to supply h do not change with respect to (2.6) and (2.7). As far as workers B are concerned, instead, conditions (2.8) and (2.9) become:

$$\mu^B(\alpha_B|h) - \mu^B(\alpha_B|i) > K-1 \quad \text{for } \theta = 1, \tag{2.11}$$

$$\frac{1}{K}\left(\mu^B(\alpha_B|h) - \mu^B(\alpha_B|i)\right) > K-1 \quad \text{for } \theta = K. \tag{2.12}$$

The same strategy as in (2.4) 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

$$\Pr(\alpha_A|h) - \Pr(\alpha_A|i) = 0.5\Pr(\theta_F = 0), \tag{2.13}$$

$$\mu^B(\alpha_B|h) - \mu^B(\alpha_B|i) = 0.5\mu^B(\theta_F = 0), \tag{2.14}$$

for workers A and B, respectively.

Combining (2.8), (2.9), (2.11), (2.12), (2.13) and (2.14) the conditions that make h the optimal choice can be rewritten:

$$\begin{aligned}
0.5\Pr(\theta_F = 0) &\geq K-1 \quad \text{for } \theta_A = 1,\\
\frac{1}{K}0.5\Pr(\theta_F = 0) &\geq K-1 \quad \text{for } \theta_A = K,\\
0.5\mu^B(\theta_F = 0) &\geq K-1 \quad \text{for } \theta_B = 1,\\
\frac{1}{K}0.5\mu^B(\theta_F = 0) &\geq K-1 \quad \text{for } \theta_B = K.
\end{aligned}$$

It can be shown by combining properly the inequalities above that the first part of the proposition is equivalent to impose that the condition

$$1 + 0.5\mu^B(\theta_F = 0) \leq K \leq \frac{1 + \sqrt{1 + 2\Pr(\theta_F = 0)}}{2} \tag{2.15}$$

holds.[14] Substituting for $\mu^B(\theta_F = 0) \geq \Pr(\theta_F = 0)$, i.e. imposing that minority workers do not overestimate the fraction of discriminatory employer, violates (2.15). Hence, the fact that workers B overestimate the percentage of discriminatory employers is a necessary condition for these inequalities to hold.[15]

Since all the workers of population A supply h and all the workers of population B supply i, the employers' strategy (2.4) implies that only workers A are promoted (second part of the proposition). Such a strategy is in turn a best reply for the employers and it can be shown it is the only strategy that is not weakly dominated under assumptions 3 and 5. The second part of the proposition amounts to verify that, when $\mu^B(\theta_F = d) > \Pr(\theta_F = d)$, wrong beliefs are always falsified unless the employer's strategy contain both "promote worker A when $e_A^1 = h, e_B^1 = i$" and "promote worker B when $e_A^1 = i, e_B^1 = h$."

■

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 exerts h - worker B exerts 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 $\mu^B(\theta_F = d) > \Pr(\theta_F = d)$

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

[15] 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 touraments between *ex ante* equal workers.

Observe first that uniqueness refers to the vector of observables (effort levels and promotions) and not to the array of strategies and beliefs that characterizes 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 all the 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 3, 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.[16] Among the various combinations of productivity levels that can be observed, it can be shown that $e_A^1 = h, e_B^1 = i$ is the only one compatible with an employers' best response that does not falsify minority workers' wrong beliefs. Another necessary condition for the wrong beliefs not to be contradicted is that also the non-discriminatory employers promote worker A after observing $e_A^1 = h, e_B^1 = i$.

The hypotheses behind Propositions 2 and 3 differ only because of the expectations of workers B. In Proposition 2 their expectations are correct, while in Proposition 3 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.[17]

[16]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).

[17]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.

Furthermore, nothing changes from the theoretical point of view when it is assumed that there are no discriminatory employers, i.e. $\Pr(\theta_F = d) = 0$, 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 2 with what happens in Proposition 3, 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.

2.3.4 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 3 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 3. 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:[18]

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

The same happens for workers A:

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

If for workers A it is still optimal to supply h, nothing changes with respect to Proposition 3, 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.[19] It is worth noting that the price for

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

[19]It is problematic to provide a sensible translation of "big enough," since this threshold

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 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.[20]

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 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.[21] Policy tools which do not change the expectations of minorities are either ineffective or very difficult to implement.

depends on many factors and therefore it varies considerably. For instance, in a situation where $\Pr(\theta_F = 0) = 1$ (absence of 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.

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

[21] 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.

2.4 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. In particular, the model can straightforwardly be contrasted with the other contributions in the literature by specifying what are the assumptions that make them directly comparable. 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.[22]

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

2.4.1 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.

[22] 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 comparison with the workers' expectations model, rather than some sort of consensus about what has been considered more important in the literature thus far. Another reason for this choice is that many detailed surveys are already available (see Blau, Ferber and Winkler (2002) and Cain (1986) among others).

The following are the assumptions that must be imposed on the model presented in Section 2.2.1 and 2.2.2, 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.

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.

2.4.2 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 seminal contribution in the statistical discrimination literature has been proposed by Arrow (1973).[23] Employer's beliefs about the existence of different characteristics between (*ex ante* identical) groups turn out to be correct in equilibrium.[24] Why are these expectations confirmed in equilibrium? In other words, why are these wrong beliefs self-confirming? The mechanism is the following: a worker's a priori unobservable variable (e.g. effort) is endogenously affected by employer's beliefs (e.g. via lower wages, or via worse job assignments), leading to a suboptimal investment in 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 2.2.1 and 2.2.2, provided that appropriate changes are made to the structure of the constituent game.[25] 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 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 with α_A standing for "assign worker A to the good job and worker B to the bad job" and vice versa for α_B.[26] Workers' strategies contain two effort levels, one if assigned

[23]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 model of workers' expectation presented in this paper and have been addressed in section 2.3.4.

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

[25]One could certainly argue that, strictly speaking, Arrow's model cannot be nested in the model of section 2.2.1 and 2.2.2, given that it is necessary to define a different game.

[26]The good and the bad job coincide with the job assigned in section 2.2.1 to the promoted

to the good job and another if assigned to the bad job. Therefore, workers' payoffs become:

$$\begin{aligned} U^{\theta_A} &= w_A - I(\alpha_A)\frac{\theta_A}{K}(e_A|\alpha_A)^2 - I(\alpha_B)\theta_A(e_A|\alpha_B)^2, \\ U^{\theta_B} &= w_B - I(\alpha_A)\theta_B(e_B|\alpha_A)^2 - I(\alpha_B)\frac{\theta_B}{K}(e_B|\alpha_B)^2. \end{aligned}$$

2. Discriminatory tastes do not play any role, i.e. the set of employer's types is a singleton $\Theta^F = \{0\}$. The employer's action is observed, and this implies that workers' expectations do not matter any more, because when they move their uncertainty has already been resolved.[27]

3. Employers believe that minority workers are on average less productive in the good job. Defining as $\mu^F(e_A|\alpha_A)$ the employer's beliefs about effort 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

$$\sum_{e_A \in E} \mu^F(e_A|\alpha_A)\, e_A > \sum_{e_B \in E} \mu^F(e_B|\alpha_B)\, e_B.$$

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 work-

and to the non-promoted worker, respectively.

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

ers' 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 chapter 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 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 2.3, 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.

2.4.3 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 in human capital, because they expect a lower lifetime return 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, mi-

nority workers find it optimal to supply on average a lower effort. Hence, they are less productive.

Some economists have heavily criticized this approach (see the next subsection), because this seemingly "voluntary" decision could actually be induced by discrimination, entering the definition of cumulative discrimination.

2.4.4 Feedback effects

The boundaries of this approach are particularly uncertain,[28] and usually surveys concerning discrimination in the labor market use these models as a counterpart 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 2.2.1 and 2.2.2, because the presence of workers' expectation is *per se* a way of formalizing such feedback effects.

[28] 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 (1973) 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 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).

2.4.5 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 caused different learning paths and different beliefs. This is a sufficient condition to observe lasting unequal outcomes in equilibrium for the two groups, even once preferences become equal, meaning that past circumstances continue to exert an influence and that expectations can be self-fulfilling.

Similarities of this chapter 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 2.2.1 and 2.2.2, 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 2.3 showed that when such an assumption is relaxed within a static framework, workers' expectations can still explain observed unequal outcomes.

2.4.6 Asymmetric tournaments

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. 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 chapter. Not surprisingly, the two models provide similar predicitons in some cases, e.g. that discriminatory tastes, as an example of unfair rules, affect the incentives of both worker A and B in the same way.

Within the literature on uneven tournaments, it is incidentally mentioned that unequal outcomes may arise between groups of workers that are *ex ante* equal.[29] 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.

[29] See Schotter and Weigelt (1992).

2.5 Conclusions

The goal of this chapter is to set up a model that analyzes the role of workes' expectations in explaining the unequal outcomes that characterize some minorities in the labor market. A framework is obtained where most of the contributions to the discrimination literature can be nested, and therefore directly compared.

The idea is that 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 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 explain the puzzling long-run persistence of cumulative discrimination.

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.

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. The next step, presented in Chapter 3, is a laboratory experiment which can provide empirical evidence about the importance of workers' expectations as a source of unequal outcomes.

Bibliography

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

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

[18] ——— (1998): *The Theory of Learning in Games.* Cambridge and London: MIT Press.

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

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

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

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

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

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

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

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

Chapter 3

Discrimination and Workers' Expectations: Experimental Evidence

3.1 Introduction

The goal of this chapter is to provide experimental evidence concerning the role of workers' expectations as an explanation for the puzzling long-run persistence of observed discrimination against certain minorities in the labor market. The model used as a benchmark (see Chapter 2) shows that *ex ante* identical groups of workers may be characterized by unequal outcomes in equilibrium due to their different beliefs. In particular, the model shows that unequal outcomes may arise when minority workers wrongly believe that they are discriminated against, even when employers do not do so either directly or statistically.

The setting of the experiment is not a carbon copy of the model. First, the experiment is designed in a way that does not remind subjects that it deals with discrimination in the labor market. This is to keep their attitudes, experiences, and opinions on the matter out of the lab, preserving the importance of payoffs in driving their behavior. The participants are randomly assigned to two populations: red and blue. In every lottery there are only two participants, one from each population. In every trial every participant has an endowment of 10 Euro cents and bids in order to win a prize worth 25 Euro cents. The players know that they face only one opponent in every trial. They also know that the higher bid wins the prize, unless they face a "crazy computer" that awards the prize to the Red regardless of the bids. Bids are not given back to either the winner or the loser. Colors (red and blue) are the equivalent of gender (or race, etc.), the prize stands for the promotion and the "crazy computers" play the role of discriminatory employers. Finally, the amount bet plays the role of effort. In the model, workers can choose to give up some utility in the first period exerting an inefficiently high level of effort in order to increase the probability of being promoted. Similarly, in the experiment subjects can trade (part of) their endowment with a higher probability of winning the prize.

This chapter analyzes whether common past experience works as a coordinating force. In other words, starting from a situation where there are discriminatory employers, do minority workers still expect being discriminated against when they do not know that discriminatory tastes have disappeared? A positive answer to this question would provide evidence in favor of the model in Chapter 2, and in particular that historical factors are crucial in selecting one among different possible outcomes (path dependent equilibrium selection), pointing toward the existence of hysteresis. This would provide useful insights concerning the long-run persistence of discrimination in the labor market.

In the lab, a reduction of bids made by disadvantaged subjects leading to a lower fraction of prizes awarded to them, is observed in three out of seven sessions, in line with the theoretical predictions, but it vanishes rather quickly during the treatment, failing to generate a Self-Confirming Equilibrium (henceforth: SCE) driven by wrong beliefs.[1] The parallel of this finding in the labor market would be a situation in which minority workers, after having been discriminated against, expect that unfavorable conditions continue while biased employers have actually disappeared. Hence, minority workers reduce their effort, and accordingly they are promoted less frequently, but eventually they discover that biased employers have disappeared and balanced promotions across populations of workers are observed.

The strategy method, on the other hand, supports the SCE driven by wrong beliefs.[2] In fact, advantaged and disadvantaged subjects react in a different way when *ad hoc* aggregate statistics are displayed. Subjects are asked to bid five times, after five different fictitious distributions of prize winners across populations in the previous period have been displayed. This fictitious distribution shows a fraction of prizes awarded to the Blues decreasing from 80% to 0% in subsequent trials. Although all subjects are informed that there is no computer that awards the prize according to the color label, what happens is that blue players are influenced by the aggregate statistics showing that a decreasing fraction of them gets the prize, and thus bid less and less. In turn, their lower bids make them less likely to win, leading to unequal outcomes, which are consistent with wrong expectations that they were less likely to get the prize.

The structure of the chapter is the following. Section 3.2 describes the theoretical framework behind the experiment and summarizes the testable implications.Section 3.3 displays the design of the experiment as well as its procedure. Section 3.4 contains the results. Section 3.5 outlines the contributions to the literature that are related to this experiment, and Section 3.6 draws some conclusions.

[1]This chapter follows Davis and Holt (1993) and Roth (1994) using the following terms:
experiment: the collection of all data;
session: the collection of data involving the same group of subjects on the same day;
treatment: a unique configuration of parameters, variables and rules;
trial: a decision unit, one repetition of the game.

[2]The strategy method, developed by Selten (1967) is a procedure that asks a group of subjects to design their strategies after having repeatedly played a game.

3.2 Theoretical framework

This chapter aims to provide experimental evidence concerning some testable implications derived from the model that analyzes the role of workers' expectations in explaining observed unequal outcomes in the labor market (see Chapter 2). This section provides a summary of the model, emphasizing its testable implications. After the experiment is presented in section 3.3, several features are contrasted and compared in more detail with the corresponding parts of the model.

The model is formalized as a two-stage game of incomplete information in which populations of workers and employers are involved. In every constituent game, i.e. in every repetition of the game played by agents randomly drawn from their populations, one employer and two workers, one of whom is a minority worker, are randomly matched. The workers choose among three levels of effort (low, intermediate, high) and the employer promotes one and only one of the two workers after having observed their effort, which is a function of observable effort and unobservable taste for work. The promotion is desirable because the job assigned to the promoted worker is assumed to be characterized by a lower cost of effort. Crucially, promotions depend via effort on workers' expectations about the unknown employer's type, which captures his possible disutility of promoting a minority worker.[3] Workers know that there are two types of employer, but they do not know whether the employer they face is discriminatory or not. Also the distribution of types within the population of employers is unknown and workers have beliefs about that.

The importance of workers' expectations can be appreciated by comparing the equilibrium outcome in terms of promotions that may arise when minority workers overestimate the percentage of employers characterized by tastes for discrimina-

[3]Observable effort and incomplete information are the main features that distinguish this approach from the tournament literature started by Lazear and Rosen (1981). The two approaches share most of their predictions, one of which being that discrimination, when it is common knowledge, affects the two populations of workers in the same way (see also Section 3.3).

tion with a situation in which their beliefs are correct *ceteris paribus*. Assuming, for the sake of simplicity, that employers do not discriminate against minorities either directly or statistically, and that all the other sources of heterogeneity such as the distribution of ability among workers have been neutralized, unequal outcomes may still arise due to minority workers' wrong expectations. In other words, wrong beliefs about being discriminated against may be self-confirming. In this circumstance what happens is that in equilibrium minority groups, who expect being discriminated against, exert less effort on average, because of a lower expected return. This induces a lower percentage of promotions within minority workers, which in turn is consistent with their beliefs that employers are characterized by discriminatory tastes. On the other hand, when beliefs are correct symmetric outcomes are observed.

It is worth stressing that a necessary condition for such a SCE is that beliefs of majority and minority workers differ. If both groups have wrong but similar beliefs about the fraction of discriminatory employers, their behavior will also be similar, and balanced outcomes in terms of promotions should be expected. The dataset used in Chapter 4 provides interesting evidence that, although men and women share very similar expectations about the magnitude of the gender wage gap, the importance they assign to the underlying causes differs. In fact, while a larger fraction of men thinks 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.

Several implications arising from this model can be tested in the laboratory:

1. *When beliefs are correct, workers' behavior should not significantly depart from the Bayes-Nash Equilibrium of the game.* In particular:

· when it is common knowledge that there is no discrimination, i.e. when the game is like a symmetric tournament, all workers should exert an inefficiently high level of effort;

· when a known amount of discrimination affects workers' behavior, there should not be systematic differences across populations. In other words, the effort exerted by majority and minority workers should decrease in a similar way.

2. *Workers' behavior should not differ across populations even when beliefs are wrong, provided that the two populations of workers share similar beliefs.*

3. *Workers who overestimate discrimination exert a lower effort than workers characterized by correct expectations.* This is the key mechanism that might drive the labor market towards unequal outcomes even when discriminatory tastes have disappeared. In Chapter 2 a static framework is used and it is assumed that minority workers are those who might have wrong beliefs. Behind the static model there is an implicit dynamic: minority workers who have experienced direct discrimination for a long period continue to expect being discriminated against even though discriminatory tastes have disappeared (hysteresis).

4. In the model it is assumed that players have costless access to aggregate outcomes that can be used to form their beliefs. An experiment can test *to what extent subjects' behavior is affected by aggregate information*, as the literature on information cascades suggests.

3.3 The experiment

The game captures the main features of the model, presented in Chapter 2, that explores the role of workers' expectations in explaining observed unequal outcomes. The game is much simpler than the model in order to be easily played. At the same time, the subjects are not aware of the underlying economic relations being tested. Thus, keywords like discrimination, labor market, employer, worker, male and female are never used. This minimizes the risk that idiosyncrasies might enter the experiment and confound the results.

3.3.1 Sketch of the game

Participants are randomly divided into two populations: red and blue. In every trial every participant has an endowment of 10 Euro cents and can decide how many cents to bid to get a prize worth 25 Euro cents. Bets are not given back to the players, neither to the winner nor to the loser, making the game equivalent to an all-pay auction. Therefore, at the end of the trial the winner gets 25 cents plus the amount not bet, while the loser only the amount not bet.

In every lottery there are only two participants, one from the red population and one from the blue population. The players know that they face only one opponent within every trial. Subjects are warned that it is possible to face the same opponent more than once during every session, but of course they do not know when, given that random assignment takes place at the beginning of each trial. The prize is awarded to the higher bid and it is split if bids are equal, unless the opponents are assigned to a "crazy computer," which instead assigns the prize to the red player regardless of the bids. The fraction of crazy computers, and whether workers know it or not, vary across treatments (see section 3.3.3).

3.3.2 Contrast and comparison with the model

Similarities with the model in Chapter 2 are straightforward only if one knows that this model is what the experiment aims to test. Colors (red and blue) are the equivalent of gender (or race, etc.). The endowment of 10 cents is the same as the utility level when intermediate effort, i.e. the optimal level of effort when promotions are not an issue, is exerted. The amount bet plays the role of additional effort exerted to enhance the probability of being promoted. The prize stands for the promotion and, finally, the "crazy computers" play the role of discriminatory employers. The game is played under different parameter settings and information structures (see section 3.3.3).

Populations and number of types

As already mentioned, red and blue labels are the equivalent of the payoff-irrelevant observable characteristic that distinguishes minority from majority workers. The color label is assigned randomly to every participant and lasts for the whole experiment.

The role of the population of employers is played by the computers, which implements the employers' equilibrium strategies in the model of Chapter 2. The crazy computers never assign the prize to the members of the blue population. The "fair computers" instead assign the prize to the player who made the higher offer and they split the prize when bids are equal. Hence, only the blue players risk being discriminated against.

In the theoretical model it is necessary to assume that workers are of different types for the employers to have some uncertainty about their productivity in the second period. In the experiment the distinction of different types would make the game much more complicated, given that subjects are not familiar with the concept of payoff-type. A further appreciable gain in simplicity is that, since the computers directly play the equilibrium strategy of the employers, it is unnecessary to play the second stage of the theoretical model.

Utility function and Nash Equilibria

The utility function used in the model is not implemented directly in the experiment because it would be quite cumbersome to deal with it in the limited time-spell of the experiment (about 75 minutes). However, the game sketched in section 3.3.1 implies a simplified but very close version of it. In both cases players have the opportunity to give up some utility with certainty in exchange for an uncertain but higher return. In the model, exerting a high effort is a sub-optimal decision considering the instantaneous utility function in the first period, but the loss of utility can be more than counterbalanced if the worker is promoted, since the job assigned

Table 3.1:
Payoffs without crazy computers (Euro cents)

red;blue	0	1	2	...	9	10
0	22.5; 22.5	10; 34	10; 33	...	10; 26	10; 25
1	34; 10	21.5; 21.5	9; 33	...	9; 26	9; 25
2	33; 10	33; 9	20.5; 20.5	...	8; 26	8; 25
...	...	...	...	...	...	...
9	26; 10	26; 9	26; 8	...	13.5; 13.5	1; 25
10	25; 10	25; 9	25; 8	...	25; 1	12.5; 12.5

to the promoted workers is characterized by a lower cost of effort. On the other hand, if the worker is not promoted she suffers a net loss of utility compared to the choice of the "safe" option, i.e. intermediate effort. The risk of not being promoted has its counterpart in the possibility to bid without getting the prize in the experiment. However, also in this case the player has a "safe" option, which is bidding zero.

Table 3.1 summarizes the payoffs for the two players given the decision rule of a fair computer who assigns the prize to the subject bidding more and splits the prize when the bids are equal. For instance, the cell 22.5; 22.5 is associated with both subjects bidding zero and the prize being split. The only Nash equilibrium occurs when both players bid 10 cents. Hence, full dissipation of the endowments is predicted, not surprisingly given that this game is equivalent to an all-pay auction with the value of the prize exceeding the sum of the endowments.

This Nash equilibrium is not efficient, because both subjects would be better off playing 0, but in that case both would have the incentive to deviate bidding a positive quantity. This is similar to what happens in the theoretical model for the combination of parameters that displays a Bayes-Nash Equilibrium with both workers supplying a high effort. Given that the model imposes that one and only one worker is promoted, if both workers offered an intermediate effort, the probability of being promoted would be the same and both would have a higher utility in the first period. However, neither worker would be maximizing her utility

Table 3.2:
Payoffs when facing a crazy computer (Euro cents)

red;blue	0	1	2	...	9	10
0	35; 10	35; 9	35; 8	...	35; 1	35; 0
1	34; 10	34; 9	34; 8	...	34; 1	34; 0
2	33; 10	33; 9	33; 8	...	33; 1	33; 0
...	...	...	...	...	...	...
9	26; 10	26; 9	26; 8	...	26; 1	26; 0
10	25; 10	25; 9	25; 8	...	25; 1	25; 0

because she has a profitable deviation supplying a high effort.

Table 3.2 summarizes the payoffs for the two players given the decision rule of a crazy computer who always assigns the prize to the red player regardless of bids. Playing 0,0 is a dominant strategy for both, just as intermediate effort is a dominant strategy in the model when there are discriminatory employers only.

When there are both crazy and fair computers, payoffs are a linear combination of these two matrices, using as weight the fraction of crazy computers. When this fraction is unknown the equivalent of the expectation-driven SCE characterized by unequal outcomes may arise, as long as crazy computers have disappeared but Blues still believe there are some.

An important difference with respect to the model is that in the experiment bids are roughly continuous, while in the model the set of choices is restricted to low, intermediate and high effort. A roughly continuous set of choices is likely to reduce the probability of observing the SCE driven by wrong beliefs, since the cost of experimenting to discover the true state of nature becomes very low (see also section 3.4).

3.3.3 Design and procedure of the experiment

One aspect of the experimental procedure needs to be stressed. All the treatments were proposed within each of the seven sessions of the experiment. Hence, all the subjects played facing the whole set of parameters. This procedure implies potential carryover effects from one parameter set to the others, as well as confounding factors arising because of framing, learning and fatigue. However, testing the existence of carryover effects (hysteresis) is one of the primary goals of this chapter, and therefore such an approach is necessary. Moreover, an econometric approach to the analysis of the data allows us to control for any observable and unobservable individual characteristic that might affect the choices of the participants during the experiment, including framing, learning and fatigue. To minimize the role of confounding factors, simultaneous parameter changes are avoided.

One of the testable implications concerned the role of information. For this reason in two sessions (3 and 4) subjects get information on all the outcomes within the session, in addition to their own outcomes, when the fraction of crazy computers is unknown (Treatments 2 and 6).

What follows is a sketch of the rules and the procedure of the experiment, which has been run using the zTree software.[4] We recruited subjects from undergraduate courses at the University of Milan. Most of the subjects were inexperienced. Participants were first randomly assigned numbers and seats. Subjects were told that their physical identity was not associated with their choices during the experiment, the subjects' numbers being their personal identification. They were given written instructions that were also read aloud by the experimenters, stressing that the amount they earned was a function of their decisions. In addition, instructions were also displayed on the screen at the beginning of each treatment.

Quiz 1. After questions are raised by subjects, a quiz is run to test their comprehension of the game. Given three different pairs of bids, they are asked to answer who is awarded the prize and to compute their earnings. If wrong answers are given, the subsequent screen shows the subject the correct answer and, in the

[4]The zTree software was developed at the University of Zurich, Institute for the Empirical Research in Economics (see Fischbacher, 2002).

case of earnings, the way to compute them. Subjects are invited again to ask questions about anything that is not clear.

Assignment to red or blue population. The color of the population is then randomly assigned to every participant by means of an algorithm, in such a way that unobserved and uncontrolled characteristics are not correlated with the focus variables. The color is assigned once and for the whole experiment.

Treatment 1 - 10 trials - Random matching of each Blue with a Red, with the possibility of facing the same opponent more than once. There is no "crazy computer" and players know it, i.e. the game is equivalent to an all-pay auction. Each subject is asked to bid from 0 to 10 Euro cents in order to get the prize. The goal of this treatment is twofold. On the one hand, it makes subjects familiar with the game. On the other hand, it tests whether players cluster around the unique Nash Equilibrium with both players bidding 10 cents, i.e. fully dissipating their endowment. Payoffs are reported in Table 3.1. After participants have decided, the computer displays:

a) how much the two opponents bid;

b) who wins the prize;

c) individual earnings.

Introduction of crazy computers. After Treatment 1, the participants are told that some crazy computers, i.e. computers that assign the prize to the member of the red population regardless of the amounts bid, will be introduced into the game. During each treatment they might face both fair and crazy computers, but the type of the computer is never known in advance. It is made clear that in every trial there is a random matching with an opponent of the other population as well as with a PC. Hence, the probability of facing a crazy computer is the same in every trial of each treatment and does not depend on the type of computer faced in previous trials.

Participants are warned that the maximum attention has been paid in order that every subject has the same expected reward. In particular, members of the blue population will receive an additional lump sum reward of 3.5 Euro to compensate them for their lower chances of being awarded the prize. Hence, at the end of the experiment everybody will have had the same chance to earn the same expected reward with differences depending on participants' actions only.[5]

Quiz 2. Subjects are then asked to answer another short quiz to test their comprehension of the game when crazy computers are introduced. Participants are asked:

a) to answer whether they are assigned the prize or not, and to compute their net position under four different circumstances in terms of bids or type of computer;

b) to compute their earnings and to infer the type of computer given four different pairs of bids and the associated decision about the prize.

In both cases, if wrong answers are given, the subsequent screen shows the subject the correct answer and, in the case of earnings, the way to compute them. Subjects are invited again to ask questions about anything that is not clear.

Treatment 2 - 15 trials - Random matching of each Blue with a Red, with the possibility of facing the same opponent more than once. There is no crazy computer but in this case players do not know this. They are just told that the fraction of crazy computers can range between 0% and 100% and that it is held constant during the treatment. Each subject is asked to bid in order to get the prize. In this case theoretical predictions are uncertain, since the best replies depend on beliefs and rationalizability does not allow us to delete any strategy. After each participant has decided the computer displays:

a) how much the two opponents bid;

[5] The lump sum compensation has been introduced to prevent members of the blue population from feeling tempted to hinder the experiment. The amount of the compensation (3.5 Euro) was derived from the difference between the expected average earnings of Reds and Blues, had they followed the theoretical predictions. It turned out that the compensation was higher than the actual difference of earnings arisen during the experiment between Reds (11.05 Euro) and Blues (9.08 Euro).

b) who wins the prize;

c) individual earnings.

In sessions 3 and 4 subjects access aggregate information as well. They are shown the distribution of bids among red and blue players, together with the percentage of Blues and Reds to whom the prize is assigned.

After statistics have been displayed, subjects are finally asked to report their beliefs about the actual fraction of crazy computers. Expectations are elicited implementing a lottery in which each subject has a probability of winning that is correlated with the number of times in which her beliefs are approximately (error $\leq 5\%$) correct.

This treatment is used as bechmark to check how subjects plays when they do not know the fraction of crazy computers and without having experienced discrimination. The same treatment is repeated after players experience discrimination (see Treatment 6). A significantly different behavior would point toward a persistent effect of observed discrimination.

Treatment 3 - 5 trials - the same as Treatment 1 except that now there is a 10% probability of facing a crazy computer and players know it. The goal of this treatment is twofold. On the one hand, it makes subjects familiar with the game when they know that crazy computers are introduced. On the other hand, it tests whether players over-react to the introduction of crazy computers, given that the unique Nash Equilibrium of the game predicts that both players should still bid 10 cents, i.e. fully dissipating their endowment (expected payoffs are reported in Table 3.3). After each participant has decided, the computer displays the same information to all the players in all sessions, i.e. in this case even in sessions 3 and 4 no one has access to aggregate information but only to individual statistics, and in particular:

a) how much the opponents offered;

b) who wins the prize;

Table 3.3:
Players' payoffs with 10 percent crazy computers (Euro cents)

red;blue	0	1	2	...	9	10
0	23.75; 21.25	12.5; 31.5	12.5; 30.5	...	12.5; 23.5	12.25; 22.5
1	34; 10	22.75; 20.25	11.5; 30.5	...	11.5; 23.5	11.25; 22.5
2	33; 10	33; 9	21.75; 19.25	...	10.5; 23.5	10.25; 22.5
...	...	...	...	...	...	...
9	26; 10	26; 9	26; 8	...	14.75; 12.25	3.5; 22.5
10	25; 10	25; 9	25; 8	...	25; 1	13.75; 11.25

c) individual earnings.

Treatment 4 - 5 trials - like Treatment 3 except that now there is a 50% probability of facing a crazy computer and players know it. As Table 3.4 shows, in this case there is no Nash Equilibrium in pure strategies, because a cycling pattern would emerge.[6] Starting from the Nash Equilibrium of the previous treatment (10;10), blue players no longer find it convenient to make a positive bid, and therefore they prefer to drop out by bidding zero. This cannot be an equilibrium because red players could get the full prize by bidding 1 instead of 10, but at that point blue players would have a profitable deviation overbidding by one cent the opponent. The same would apply to red players and so on, until the mechanism starts again at (10,10). There is only one prediction: the average bid of Reds and Blues should not significantly differ. After participants decide, the computer displays the same individual statistics as in the previous treatment.

Treatment 5 - 5 trials - equal to Treatment 4 except that now there is a 90% probability of facing a crazy computer and players know it. As Table 3.5 shows,

[6]A mixed strategy equilibrium has been computed using a continuous support $[0, 10]$ for the strategy space:
- Reds assign a uniform distribution to the whole support $g(s) = 2/25, s \in [0, 10]$ and in addition $p(s = 0) = 1/5$;
- Blues assign a uniform distribution to the whole support $g(s) = 2/25, s \in [0, 10]$ and in addition $p(s = 10) = 1/5$.

Table 3.4:
Players' payoffs with 50 percent crazy computers (Euro cents)

red;blue	0	1	2	...	9	10
0	28.75; 16.25	22.5; 21.5	22.5; 20.5	...	22.5; 13.5	22.25; 12.5
1	34; 10	27.75; 15.25	21.5; 20.5	...	21.5; 13.5	21.25; 12.5
2	33; 10	33; 9	26.75; 14.25	...	20.5; 13.5	20.25; 12.5
...	...	...	...	...	...	...
9	26; 10	26; 9	26; 8	...	19.75; 7.25	13.5; 12.5
10	25; 10	25; 9	25; 8	...	25; 1	18.75; 6.25

Table 3.5:
Players' payoffs with 90 percent crazy computers (Euro cents)

red;blue	0	1	2	...	9	10
0	33.75; 11.25	32.5; 11.5	32.5; 10.5	...	32.5; 3.5	32.25; 2.5
	34; 10	32.75; 20.25	31.5; 10.5	...	31.5; 3.5	31.25; 2.5
	33; 10	33; 9	31.75; 9.25	...	30.5; 3.5	30.25; 2.5
		...	...	...	...	...
		26; 9	26; 8	...	24.75; 2.25	23.5; 2.5
		25; 9	25; 8	...	25; 1	23.75; 1.25

although there is no Nash Equilibrium in pure strategies, it is possible to predict that both players should bid from 0 to 2 cents, because bidding more than 2 cents is never a best reply to any beliefs. Moreover, a mixed strategy equilibrium exists when both red and blue subjects assign the following probabilities: $p(s = 0) = 1/5; p(s = 1) = 3/5; p(s = 2) = 1/5$. After the participants have decided, the computer displays the usual individual statistics.

Treatment 6 - 15 trials - Repetition of Treatment 2: random matching of each Blue with a Red, with the possibility of facing the same opponent more than once. There is no crazy computer but players do not know it; they are just told that the fraction of crazy computers can range between 0% and 100% and that it is held constant during the treatment. As in Treatment 2, individual statistics are displayed after the participants have decided, while in the same two sessions (3 and 4) subjects access aggregate information as well. This is identified as a different treatment because of carryover effects that might affect subjects' behavior. Average bids that significantly differ from Treatment 2 would signal persistent effects of discrimination. Blues bidding less than Reds would be evidence in favor of the SCE in the model.

Since the set of choices in the experiment is roughly continuous, the SCE that may arise differs from the SCE in the model. As in the model, minority workers should choose the "safe" option, i.e. bidding zero (equivalent of intermediate effort), as long as they think there is a sufficiently high fraction of crazy computers. However, while in the model the majority workers would exert high effort, in the game it is enough that Reds bid 1 cent instead of 10 (which would be the equivalent of high effort) if they believe that the Blues play zero. In more detail, the SCE driven by wrong minority workers' beliefs corresponds in the game to a situation in which:

- Reds bid 1 cent, which is a best reply to beliefs that crazy computers are up to 8% and that Blues bid zero.
- Blues think that at least 92% of the computers are crazy and that Reds bid 1 cent. Given this beliefs bidding zero is a best reply.
- The prize is awarded to the Reds.[7]

Treatment 7 - 5 trials - Repetition of Treatment 3: 10% probability of facing a crazy computer and players know it. Again, this is identified as a separate

[7]This SCEquilibrium is not unique. Another equilibrium is obtained simply flipping the color labels in the example above. However, only the equilibrium in which the prize is assigned to the Reds is consistent with the model. If the equilibrium in which the prize is assigned to the Blues occurs, it would be strong evidence against the model.

On the other hand, if all players believe that at least 92% of the computers are crazy and that the opponent play zero, both Reds and Blues should bid zero. However, in this case a few trials are enough to falsify players beliefs that there are many crazy computers, given that the prize would be split every time.

Finally, another Self-Confirming Equilibrium is supported by both subjects bidding the whole endowment and thinking that there are at most 20% of crazy computers.

Table 3.6:
Summary of rational play as predicted by pure strategy NE

Treatment	Reds	Blues
1	10	10
2	?	?
3	10	10
4	?	?
5	≤ 2	≤ 2
6	?	?
7	10	10

treatment because of carryover effects that might affect subjects' behavior. Significantly different bids than in Treatment 3 would point toward persistent effects of discrimination. Blues bidding less than Reds would be evidence in favor of the SCE driven by wrong beliefs.

Table 3.6 summarizes the rational play, i.e. how many Euro cents subjects should bid, according to the Nash Equilibria in pure strategies.

Strategy method. - 5 trials - There is no crazy computer and players know it. Each subject is asked to bid five times, every time after a different fictitious distribution of prizes between populations in the previous round has been displayed. This fictitious distribution shows a fraction of prizes won by the Blues decreasing from 80% to 0% in subsequent periods. In this case each subject wins the prize if her bid exceeds the average bid of the population of opponents. The individual results of each trial are shown together after the fifth trial, so that they do not affect the choices of the subjects during the treatment. Since there is no crazy computer, and subjects know it, subjects of both populations should dissipate entirely their endowment and pay no attention to the aggregate statistics.

At the end of the experiment a questionnaire is proposed, reminding participants that their physical identity was not associated with their choices and their answers during the experiment. Questions concerned academic as well as personal information. In section 3.3.4 some descriptive statistics of the pool of subjects are summarized.

70 subjects participated in the experiment, which consisted of seven sessions. The sessions lasted approximately 75 minutes and were composed of a minimum of 8 and a maximum of 12 subjects. Euro cents were the currency used during the experiment. Earnings ranged between 9 and 15.5 Euros (11.8 on average). Average earnings of the Reds were lower than average earnings of the Blues once the lump sum compensation of 3.5 Euros was taken into account. (11.05 Euro vs. 9.08+3.5 Euro, respectively).

3.3.4 Sample description

From the information collected by means of the final questionnaire, it turns out that males are over-represented in our sample (67% vs. 33%), and that the average age of the pool is about 21 years. Most of the participants (89%) comes from the School of Political Sciences, and is enrolled in the third year of the degree program. The final mark at the exit of secondary school was chosen as a proxy for a student's ability; the variable has been rescaled in the range [0,1]. Two thirds of the sample come from high schools (*licei*) and one fourth from technical schools (*istituti tecnici*). Two specific questions were asked concerning political and religious orientation. An ordered scale from 0 to 5 has been used to ask subjects about their political orientation (0=left; 5=right), without any label on each possible choice. Two thirds of the subjects report themselves as being center-left, i.e. they chose a value from 0 to 2, and 33% center-right. The average choice is 1.97 while the median choice is 2. With respect to religion, the subjects have been asked to choose from three alternatives: "believer and churchgoer," "believer but not churchgoer," "non-believer." The proportion of the last occurrence was around one third.

3.4 Results

From the quizzes it is possible to infer that subjects have a good albeit imperfect comprehension of the game, given that the average number of wrong answers is about 1.4 out of 6 questions in the first quiz, and 2.5 out of 16 in the quiz with crazy computers.

The results confirm the prediction implied by the Nash equilibrium of the game where subjects know that there is no crazy computer (**Treatment 1**), i.e. that they should fully dissipate their endowment and bid 10 cents. In fact, the average bid is 9.14 cents. In the last trial of Treatment 1, 64 out of 70 subjects fully dissipate their endowment. The distribution of prizes is balanced across populations, in line with predictions. The prize is split in 63.7% of the games, while Reds win 16.6% of the games and Blues 19.7%. In the last trial of the Treatment the prize is split in 30 out of 35 games. Table 3.7 reports the average bid of each population, session by session, pooling all periods. The last row summarizes the same statistics pooling all the sessions. Blues display a propensity to bid more than Reds in session 5, although the difference is not statistically significant at the 10% confidence level. In fact, the Kolmogorov-Smirnov test does not reject the equality of the distribution functions (p:0.112). In the other sessions, as well as in the pooled sample, the behavior of red and blue subjects is very similar, in line with the theoretical predictions.

Let's skip for one moment the analysis of Treatment 2. When a known and small fraction (10%) of crazy computers is introduced (**Treatment 3**), some subjects slightly over-react (see Table 3.8). In fact, while the small fraction of crazy computers does not affect the prediction of full dissipation of the endowment, the average bid decreases to 8.3 cents with a higher variability. This is particularly evident in session 3 and, to a minor extent, in sessions 2, 4, 6 and 7. Reds and Blues display a pattern that is not significantly different in each session or in the pooled sample, in line with predictions. Unequal outcomes emerge as a consequence of

Table 3.7:
Treatment 1: average bid (st dev) and distribution of prize winners

Session (Subjects)		Average bid red	Average bid blue	Kolmogorov-Smirnov test	Prize winners red	Prize winners split	Prize winners blue
1	(10)	9.82(0.66)	9.70(0.81)	p:1.000	6	41	3
2	(10)	9.58(1.39)	9.76(0.72)	p:1.000	5	37	8
3	(10)	9.02(2.80)	9.56(1.51)	p:0.998	8	37	5
4	(8)	8.95(1.28)	8.57(2.51)	p:0.579	11	14	15
5	(10)	7.44(3.92)	8.92(2.55)	p:0.112	9	23	18
6	(10)	9.80(1.28)	9.76(0.85)	p:1.000	5	43	2
7	(12)	8.62(2.73)	8.63(2.76)	p:0.928	14	28	18
pooled	(70)	**9.02(2.42)**	**9.27(1.93)**	p:**0.942**	**58**	**223**	**69**

Table 3.8:
Treatment 3: average bid (st dev) and distribution of prize winners

Session (Subjects)		Average bid red	Average bid blue	Kolmogorov-Smirnov test	Prize winners red	Prize winners split	Prize winners blue
1	(10)	9.60(0.66)	9.88(0.60)	p:1.000	2	22	1
2	(10)	7.20(4.20)	9.32(2.08)	p:0.285	7	11	7
3	(10)	7.16(4.33)	6.40(4.25)	p:0.710	11	7	7
4	(8)	8.45(2.43)	7.75(3.93)	p:0.832	7	5	8
5	(10)	7.88(3.26)	8.76(2.17)	p:0.915	9	8	8
6	(10)	9.60(1.38)	8.04(3.79)	p:0.915	5	19	1
7	(12)	8.27(3.49)	7.83(3.58)	p:0.958	13	12	5
pooled	(70)	**8.30(3.29)**	**8.29(3.28)**	p:**1.000**	**54**	**84**	**37**

Table 3.9:
Treatment 4: average bid (st dev) and distribution of prize winners

Session (Subjects)		Average bid red	Average bid blue	Kolmogorov-Smirnov test	Prize winners red	Prize winners split	Prize winners blue
1	(10)	4.80(4.72)	3.72(4.42)	p:0.475	17	5	3
2	(10)	5.20(4.06)	4.32(4.12)	p:0.475	22	2	1
3	(10)	1.12(1.39)	5.96(3.77)	p:0.000	13	0	12
4	(8)	2.65(3.72)	4.40(3.91)	p:0.034	15	2	3
5	(10)	4.28(3.32)	4.72(4.01)	p:0.710	22	1	2
6	(10)	6.76(4.32)	5.84(4.52)	p:0.915	16	6	3
7	(12)	3.80(4.02)	4.77(3.94)	p:0.071	18	3	9
pooled	(70)	**4.12(4.09)**	**4.83(4.11)**	p:**0.313**	**123**	**19**	**33**

the crazy computers, with Blues winning only 21.1% of the prizes against 30.8% of the Reds.

When the fraction of crazy computers increases to 50% in **Treatment 4**, the average bid decreases sharply to 4.47 cents (4.12 the Reds, 4.83 the Blues), with Blues bidding clearly more than Reds in most of the sessions. The Kolmogorov-Smirnov test rejects the equality of distribution functions in session 3 (p=0.000), 4 (p=0.034) and 7 (p=0.071), while the difference is not statistically significant in the other sessions or in the pooled sample (see Table 3.9). The effect of the high fraction of crazy computers is overwhelmingly stronger than the effect of the higher offers of blue subjects in determining the distribution of prize winners (70.2% Reds Vs 18.8% Blues).

When in **Treatment 5** 90% of computers are crazy, the average bid (1.52 cents) stays within the predicted range of [0,2] cents, but there is a lot of variability across sessions. As Table 3.10 shows, red subjects bid on average a quantity that cannot be a best reply to any beliefs in sessions 2, 6 and 7. Reds bid significantly more than Blues in session 2, while the opposite happens in sessions 3 and 4. The difference is not statistically significant in the other sessions or in the pooled sample (p=0.203). Unequal outcomes become particularly severe, with Reds winning the prize 93.7% of the times, against 4.6% of the Blues.

Table 3.10:
Treatment 5: average bid (st dev) and distribution of prize winners

Session (Subjects)		Average bid red	Average bid blue	Kolmogorov-Smirnov test	Prize winners red	Prize winners split	Prize winners blue
1	(10)	0.96(1.65)	0.92(2.16)	*p*:0.915	24	0	1
2	(10)	3.88(3.55)	1.28(2.11)	*p*:0.015	24	0	1
3	(10)	0.36(0.56)	2.00(2.19)	*p*:0.002	23	2	0
4	(8)	0.15(0.67)	1.85(2.87)	*p*:0.000	17	0	3
5	(10)	0.72(0.68)	0.80(1.44)	*p*:0.996	24	1	0
6	(10)	3.08(4.14)	1.44(2.61)	*p*:0.710	23	0	2
7	(12)	2.07(3.68)	1.53(2.10)	*p*:0.135	29	0	1
pooled	(70)	**1.66(2.94)**	**1.39(2.23)**	*p*:**0.203**	**164**	**3**	**8**

Summarizing, subjects follow rather well the theoretical predictions when the fraction of crazy computers is known. This is particularly true in session 1, where deviations from the predicted behavior are negligible. Some departures from the predicted behavior are worth noting, however. In sessions 3 and 4 Blues have the propensity to bid significantly more than Reds when a sufficiently large fraction of crazy computers is introduced. This happens also in one case in session 7. In session 2 red players react to the presence of a negligible fraction of crazy computers. Also in session 6 (blue players) and in session 3 (all players) there is evidence of over-reaction when the fraction of crazy computers is negligible. Reds offer on average an amount higher than rationalizable in sessions 2, 6 an 7 when many crazy computers are introduced.

The analysis of **Treatment 2**, the first with the fraction of crazy computers equal to zero but unknown to the players, shows a high variability both between and within sessions. This is not surprising, given that any bid can be a best reply given some beliefs about the fraction of crazy computers and about the opponent's strategy. What is interesting is the fact that disadvantaged players bid on average significantly more than the opponents in three out of seven sessions (3, 5 and 7). In session 7 this happens because only red players significantly react to the possibility that crazy computers are introduced after Treatment 1, while blue players do not.

Table 3.11:
Treatment 2 - all periods: average bid (st dev) and distribution of prize winners

Session (Subjects)		Average bid red	Average bid blue	Kolmogorov-Smirnov test	Prize winners red	Prize winners split	Prize winners blue
1	(10)	8.97(2.48)	8.07(3.90)	p:0.518	14	46	15
2	(10)	7.16(4.03)	7.72(3.31)	p:0.395	14	39	22
3	(10)	5.90(4.62)	6.81(3.62)	p:0.010	26	19	30
4	(8)	5.93(4.25)	5.01(4.22)	p:0.375	28	8	24
5	(10)	5.21(4.06)	6.67(3.50)	p:0.042	27	7	41
6	(10)	9.07(2.04)	8.69(2.68)	p:0.996	18	44	13
7	(12)	6.70(4.36)	7.78(3.56)	p:0.081	22	38	30
pooled	(70)	**7.01(4.04)**	**7.33(3.68)**	p:**0.030**	**149**	**201**	**175**

In session 5 both populations significantly reduce their bids in a similar way, but Blues offer more than Reds also in Treatment 1, when it is known that there was no crazy computer. In session 3 both populations significantly reduce their bids, but Reds more than Blues. In a nutshell, the announcement that there is the possibility of facing crazy computers significantly affects the behavior of most of the subjects, but in some cases Reds react more than Blues.[8]

The fraction of crazy computers being unknown, it is interesting to see what happens when the analysis is restricted to the last 5 periods of the treatment, when subjects have the possibility of learning from their experience (and in sessions 3 and 4 also from aggregate statistics) the true state of nature, i.e. that there is, in fact, no crazy computer. Data reported in Table 3.12 show that learning takes place. In all the sessions the average bid is higher at the end of the treatment with the exception of blue players in session 3 and red players in session 4, the two sessions where aggregate information is available. However, evidence is not strong enough to claim that players in these two cases are learning the "wrong" SCE with discrimination.

[8]The Kolmogorov-Smirnov test does not reject at 10% confidence the invariance of players' behavior between treatments 1 and 2 only in sessions 1 and 6 and only for blue players in session 7. These statistics are not reported in detail in order to save space, but are available from the author.

Table 3.12:
Treatment 2 - first 5 and last 5 periods: average bid (st dev)

Session (Subjects)		Beginning of Treatment 2 Average bid		End of Treatment 2 Average bid	
		red	blue	red	blue
1	(10)	7.88(3.01)	6.60(4.66)	9.96(0.20)	9.60(2.00)
2	(10)	6.92(4.38)	7.36(5.39)	7.72(3.77)	8.04(3.31)
3	(10)	6.28(4.70)	7.28(3.21)	7.00(4.39)	6.32(3.91)
4	(8)	6.15(4.46)	4.35(4.16)	5.60(4.33)	5.20(4.25)
5	(10)	3.08(3.29)	6.04(3.06)	6.56(3.90)	7.60(3.58)
6	(10)	8.00(2.55)	7.68(3.56)	9.80(1.00)	9.32(1.89)
7	(12)	5.73(4.33)	7.10(3.89)	7.47(4.30)	8.13(3.51)
pooled	(70)	**6.28(4.13)**	**6.70(3.82)**	**7.78(3.76)**	**7.83(3.52)**

Table 3.13:
Treatment 6 - all periods: average bid (st dev) and distribution of prize winners

Session (Subjects)		Average bid		Kolmogorov-Smirnov test	Prize winners		
		red	blue		red	split	blue
1	(10)	9.64(1.70)	9.08(2.77)	p:0.996	9	63	3
2	(10)	7.21(4.04)	9.67(1.18)	p:0.002	6	43	26
3	(10)	7.08(4.45)	6.12(3.77)	p:0.003	31	12	31
4	(8)	5.33(3.99)	5.20(3.95)	p:0.047	27	7	26
5	(10)	6.35(4.35)	8.01(3.10)	p:0.016	16	25	34
6	(10)	9.96(0.35)	9.63(1.50)	p:0.996	5	69	1
7	(12)	9.32(2.39)	8.97(2.70)	p:0.948	14	69	7
pooled	(70)	**7.96(3.69)**	**8.20(3.23)**	p:**0.050**	**108**	**299**	**118**

Let us move to the repetition of the same treatment after subjects have faced an increasing fraction of crazy computers (10%, 50% and 90%), summarized in Table 3.13. In sessions 2 and 5 Blues offer significantly more than Reds, while in session 3 Reds bid significantly more than Blues.

One of the key questions this chapter tries to answer is how subjects' behavior differs before and after having experienced the experiment's version of discriminatory tastes. In sessions 5, 6 and 7 subjects bid significantly more after having faced crazy computers than before. This also happens to Reds in session 1 and to Blues in session 2, respectively. In the other cases significant differences do not

Table 3.14:
Treat. 2, last 5 periods; Treat. 6, first 5 and last 5 periods: average bid (st dev)

	End of Treatment 2 Average bid		Beginning of Treat. 6 Average bid		End of Treatment 2 Average bid	
	red	**blue**	**red**	**blue**	**red**	**blue**
1	9.96(0.20)	9.60(2.00)	9.32(2.17)	8.00(4.08)	9.60(2.00)	9.40(2.29)
2	7.72(3.77)	8.04(3.31)	8.20(3.49)	9.32(1.84)	6.80(4.40)	10.00(0.00)
3	7.00(4.39)	6.32(3.91)	7.04(4.49)	5.20(3.62)	6.40(4.90)	6.80(3.85)
4	5.60(4.33)	5.20(4.25)	5.50(3.90)	3.75(3.49)	5.30(4.23)	6.30(4.01)
5	6.56(3.90)	7.60(3.58)	5.52(4.38)	6.96(3.51)	6.36(4.59)	8.96(2.32)
6	9.80(1.00)	9.32(1.89)	9.88(0.60)	9.20(2.00)	10.00(0.00)	9.68(1.60)
7	7.47(4.30)	8.13(3.51)	8.63(3.19)	9.00(2.46)	10.00(0.00)	8.53(3.34)
	7.78(3.76)	**7.83(3.52)**	**7.82(3.69)**	**7.50(3.49)**	**7.91(3.85)**	**8.58(3.01)**

emerge, but a significant decrease of bids is never observed. At first glance, this finding is rather puzzling given that the opposite should be expected following the predictions in Chapter 2. However, as Table 3.14 shows, restricting our attention to the beginning of **Treatment 6** as opposed to the end of treatment 2, it turns out that in sessions 1, 3 and 4 Blues reduce their bids, while Reds do not. In theory, this should lead to the SCE driven by wrong beliefs to emerge, but this effect on disadvantaged subjects' behavior is short-lived and vanishes during the treatment. In fact, at the end of the treatment red and blue subjects behave in a very similar way in most of the sessions.

Hence, evidence about hysteresis is rather weak, because the finding that, after having been discriminated against, Blues bid less than Reds and receive a lower fraction of prizes is not long-lasting. Although a SCE with wrong beliefs is something that can happen but cannot be expected to emerge regularly, evidence is not fully satisfactory if evaluated from the point of view of the testable implication 3 above. The main reason the experiment does not provide strong evidence in favor of unequal outcomes driven by subjects' expectations, is that the experiment failed to separate the beliefs of the two populations. In the theoretical model different expectations about the fraction of discriminatory employers are a necessary condition for the SCE to be observed, but in the lab Blues never expect a clearly higher fraction of crazy computers than Reds. A possible explanation for this finding is that the experimental design overemphasized the discontinuity between treatments, preventing carryover effects from entering the picture. Another possible explanation could be that the set of choices was roughly continuous. This induces a very slight difference in the optimal behavior of advantaged and disadvantaged subjects, and therefore also a very low cost of experimenting to discover the true state of nature. Were the choice set dichotomized (e.g. 0, 5 and 10 cents), it would be more costly to experiment and the Self-Confirming Equilibrium would be more likely to emerge.

From **Treatment 7**, i.e. the repetition of treatment 3 after subjects face an increasing fraction of crazy computers, subjects' behavior does not change, with the exception in session 4 of Reds who significantly reduce their bid.[9]

Quite surprisingly, given the evidence from the comparison of Treatments 2 and 6, the strategy method shows evidence supporting the Self-Confirming Equilibrium driven by wrong beliefs. In fact, advantaged and disadvantaged subjects react in a different way when *ad hoc* aggregate statistics are displayed showing a fraction of prizes won by the Blues decreasing from 80% to 0% in subsequent periods. The subjects are told that there is no crazy computer. Hence, subjects of both populations should pay no attention to the aggregate statistics and entirely dissipate their endowment. On the contrary, as the fictitious distribution of prizes becomes less and less favorable to blue players, their offers decrease. On the other hand, red players do not change their behavior significantly (see Table 3.15).

As a result, the fraction of blue players bidding more than the average of their opponents also decreases from 74.3% in the first period of the strategy method, when the distribution of prizes was supposed to be more favorable, to 25.7% in the

[9]These statistics are not reported in detail in order to save space, but are available from the author.

Table 3.15:
Strategy method: average bid (st. dev) across populations period by period

		Percentage of prizes announced to be won by Blues				
	pop	**80**	**60**	**40**	**20**	**0**
1	**red**	6.80(4.20)	8.00(2.24)	8.80(1.00)	9.00(1.58)	8.60(2.86)
1	**blue**	8.80(1.98)	8.60(2.00)	4.00(3.82)	3.40(4.39)	0.00(0.00)
2	**red**	8.80(1.98)	9.00(1.29)	7.60(3.96)	3.20(3.32)	4.20(4.85)
2	**blue**	9.40(1.22)	9.00(1.29)	6.40(3.07)	2.20(4.00)	2.00(4.85)
3	**red**	4.20(4.85)	5.20(4.58)	7.20(4.00)	7.20(4.00)	4.40(4.68)
3	**blue**	7.80(1.75)	8.00(1.71)	7.80(2.45)	6.80(4.00)	3.60(4.55)
4	**red**	7.00(4.23)	5.75(3.43)	5.00(3.16)	5.75(3.73)	4.25(4.49)
4	**blue**	8.00(2.05)	8.25(1.97)	4.50(4.26)	4.00(4.35)	3.75(4.25)
5	**red**	6.20(4.40)	8.80(2.45)	8.60(1.53)	9.40(0.82)	7.60(3.57)
5	**blue**	7.20(3.73)	10.00(0.00)	8.00(1.94)	6.00(5.00)	4.20(4.85)
6	**red**	8.00(4.08)	7.60(3.27)	9.80(0.40)	10.00(0.00)	10.00(0.00)
6	**blue**	8.20(3.67)	7.40(2.55)	7.60(3.96)	2.60(3.85)	4.20(4.85)
7	**red**	6.00(4.23)	8.17(2.65)	7.33(3.60)	9.50(0.77)	7.00(4.35)
7	**blue**	7.16(3.44)	7.50(2.33)	4.67(2.86)	3.67(4.57)	3.17(4.00)
pooled	**red**	**6.69(4.25)**	**7.57(3.22)**	**7.83(3.15)**	**7.83(3.33)**	**6.66(4.37)**
pooled	**blue**	**8.06(2.83)**	**8.37(2.04)**	**6.14(3.57)**	**4.09(4.54)**	**2.97(4.27)**

fifth period without any change in fundamentals. What happens in all sessions is that blue players are influenced by the aggregate statistics showing that less and less of them has won the prize, inducing them to bid less. In turn, their lower bids make them less likely to win, leading to unequal outcomes that are consistent with wrong expectations that they were less likely to get the prize.

Given the design of the experiment, and in particular that every subject is exposed to the whole set of parameter changes, a regression analysis of the data is certainly informative. The limitations imposed by the very low number of independent observations prevent inference from being reliable. However, the interpretation of a regression as a conditional expectation function is not at all affected by the low number of independent observations and sheds more light on the data. In a multivariate framework, with players' bids as a dependent variable, where fixed effects control for any observable or unobservable individual characteristic, session fixed effects display a high heterogeneity. The fraction of crazy computers, or individual beliefs when the fraction of crazy computers is not known, is obviously the most important variable in explaining the variation of bids, accounting for a 0.61 cents lower bid every 10% of crazy computers. It is not, however, the only one. The dummy variable for the population shows that, everything else being equal, the blue players have a propensity to bid much more (3.5 cents). As far as the learning and framing effects, bids tend to increase within treatments where the fraction of crazy computers is not announced, while across treatments a U-shaped negative effect emerges, with a minimum in treatment 4 where bids are *ceteris paribus* about 1.3 cents lower than in treatments 1 and 7.

3.5 Literature review

Although the role of workers' expectations in explaining unequal outcomes has never been the focus of experiments, several contributions to the literature are relevant as far as this experiment is concerned. They can be divided into three groups:

1. Discrimination and asymmetric tournaments.
2. Information cascades, sunspot and hysteresis.
3. All-pay auctions.

3.5.1 Experimental studies of discrimination and asymmetric tournaments

Experiments closely related to the experiment presented in this chapter are those concerning either statistical discrimination or asymmetric tournaments. This subsection concentrates on experiments based on economic factors. A survey of many other experiments based on group identification or status can be found in Anderson, Fryer and Holt (2002).

The literature concerning experimental studies of discrimination is thin and in general not directly related to the experiment presented in this chapter, with a few exceptions. Fryer, Goeree and Holt (2002b) describe the results of experiments that may produce (and sometimes do) a pattern of experience-based discrimination consistent with the statistical discrimination models proposed by Arrow (1973) and Phelps (1972). Employers have to decide whether to hire or not workers from two otherwise identical populations, "green" and "purple." The hiring decision is affected by an observable test score, which in turn depends on a worker's (unobserved) investment decision, like education or training. The cost of investing is random and it is set to be systematically higher for the workers of one population during the first ten out of sixty rounds, while from the eleventh onward it drawn from the same distribution. Moreover, players have access to aggregate information, given that the average investment and hiring percentages for the workers of each color are displayed at the end of each round. The authors find that a different average investment emerges, and then a lasting and self-reinforcing mechanism operates in such a way that fewer workers of that group are hired, the fraction of that group of workers investing decreases even further and so on, leading to multiple equilibria with discrimination.[10] There are certain dimensions in which this experiment should be explicitly compared with that presented above. In particular, it is worth noting that, similarly to the experiment described in this chapter:

a) there is a real effect (the different distributions from which investment costs are drawn) that is withdrawn during the experiment, but that have long-lasting effects (hysteresis);

b) aggregate statistics are used to convey information to players, although there is no specific treatment focusing on the presence vs. absence of aggregate information.

[10]Fryer, Goeree and Holt (2002a) present classroom experiments very similar to Fryer, Goeree and Holt (2002b), but where a clear relation going from a higher average cost to a lower propensity to invest does not emerge.

c) there is an endogenous decision (investing or not) that makes *ex ante* equal populations potentially different in equilibrium.

Hargreaves-Heap and Varoufakis (2002) use the Hawk-Dove game to show that starting from two populations that differ because of a payoff-irrelevant observable characteristic only (red and blue label), different roles associated with different payoffs (i.e. discriminatory conventions) may emerge. Löhm (2000) finds in a Battle of the Sexes experiment that females are more likely to be discriminated against by other females.[11]

The literature concerning experimental studies of asymmetric tournaments is more established and some papers can fruitfully be used as a benchmark, in particular Schotter and Weigelt (1992) and Bull, Schotter and Weigelt (1987). The former presents an experiment aiming to test the theoretical predictions of the asymmetric tournament theory as presented by O'Keeffe, Viscusi, and Zeckhauser (1984). In particular, they focus on the predicted trade-off between equity and efficiency associated with affirmative actions, finding contrary evidence.

From the theoretical point of view, asymmetric tournaments have many things in common with the model in Chapter 2. In line with the old saying that different opinions are necessary for a horse race to take place, both involve uncertainty. What distinguishes them is the fact that effort is not perfectly observable in the asymmetric tournament literature, while incomplete information about the opponents' type-strategy set characterizes Chapter 2. Furthermore, the two approaches share most of their predictions, in particular that the behavior of advantaged and disadvantaged workers should change in a similar way when discrimination is common knowledge.

[11]Other experiments concerning statistical discrimination have been proposed by Davis (1987) and Anderson and Haupert (1999), both relying on exogenous differences that characterize the two populations. The former finds weak evidence that the larger population has better outcomes. The latter provides evidence that workers belonging to a population characterized by a lower average innate productivity are less likely to be hired, with the likelihood depending on the cost of discovering the individual type. Strictly speaking, it can be argued that the framework of these experiments cannot be classified as discrimination.

The experiment presented by Schotter and Weigelt (1992) can be used as a benchmark also from the methodological point of view. Two points are particularly relevant. First, the authors want to avoid carryover effects from one treatment to another. Consequently, each subject is allowed to participate in one treatment only. In the present chapter, instead, the main goal is to test the existence of hysteresis, and therefore carryover effects are part of the picture. Hence, the treatments are designed in such a way that every subject faces both a symmetric game and situations characterized by discrimination. Second, in the experiment proposed by Schotter and Weigelt (1992) players are matched once and for all within every session. This is more likely to lead to cooperation, or at least to strategic interaction, that would instead disappear with a random matching repeated before every period (see Duffy and Ochs (2003)).

The experiment just described closely follows an earlier experiment by Bull, Schotter and Weigelt (1987), where asymmetries and affirmative actions were not the main focus. It is worth noting that both experiments report a tendency of disadvantaged workers to over-supply effort in uneven tournaments, as if asymmetries elicit greater effort.[12] Moreover, Schotter and Weigelt (1992) show that advantaged workers make more effort than predicted in unfair tournaments. Section 3.4 shows evidence that disadvantaged players sometimes over-bid in unfair contests, while some evidence emerges that advantaged players bid more than predicted when the rules of the tournament are particularly favorable.

3.5.2 All-pay auctions

All-pay auctions are characterized by the fact that bids are given back neither to the winner nor to the loser. The model behind the experiment is related to an all-pay auction, insofar as there is no compensation for the loss of utility that a non-

[12] Asymmetric contests are defined "uneven" when agents are different, and "unfair" when contestants are identical but the rules favor one of them.

promoted worker suffers when he have exerted an effort higher than the level that would be optimal if promotion was not an issue. In an all-pay auction the prize goes to the highest bidder, so that each player has the incentive to overbid the others, as long as this ensures a positive payoff.[13] When the value of the prize exceeds the sum of the endowments an equilibrium in pure strategies exists, implying full dissipation of the endowments. Otherwise, in symmetric all-pay auctions, the result that the sum of the expected bids equals the value of the prize is supported by mixed strategies equilibria (Baye, Kovenock and de Vries, 1996). Rational agents never over-dissipate the value of the rent if they have the opportunity to bid zero. However, a relaxation of the rationality via the possibility of decision errors is enough to support a theoretical framework where over-dissipation can be observed (see Anderson, Goeree and Holt, 1998) consistently with experimental evidence like Davis and Reilly (1998).[14]

The promotion game without discrimination in Chapter 2, which is also tested in this experiment, is equivalent to an all-pay auction. Not surprisingly, the Nash equilibrium is a corner solution where endowments are fully dissipated. It might be feasible to increase the value of the endowment (or to decrease the value of the prize) to test also for over-dissipation. However, this goes beyond the goal of this chapter.

[13]The literature on symmetric rank-order tournaments started by Lazear and Rosen (1981) shares some of the features concerning all-pay auctions.

[14]When within such a framework subsequent bids are allowed before the prize is assigned, it is easy to observe that bidding spirals out of control, as in the Dollar Auction Game presented by Shubik (1971), where a dollar is awarded to the highest bid. Since the expenditures are sunk, it would be rational to increase a bet whenever doing so increases the expected return more than the amount of the additional bet. There is no stable equilibrium (at least in pure strategies) as long as the endowment of each player exceeds the value of the prize. When one bid exceeds the value of the prize, the motivation of the remaining bidders changes from a desire to maximize returns to one of minimizing losses. Thus, the question transforms from "How much can I win?" to "How do I keep from losing?" and escalation is easily observed, like in the classroom experiments described by Murnigham (2001).

3.5.3 Information cascades, sunspots and hysteresis

This experiment is also related to the strand of literature on information cascades (or herd behavior) started by Banerjee (1992) and Bickchandani, Hirshleifer and Welch (1992). Information cascades occur when the initial decisions of other players coincide in a way that it is optimal for each of the subsequent individuals to ignore his or her private signals and follow the established pattern. Particularly interesting as far as this chapter is concerned is the sub-case of a reverse cascade (also called bad herd, or lemming type behavior), which happens when the first decision-makers choose the incorrect state of the world, and the followers join the resulting pattern of mistakes despite the fact that their private signals are more likely to indicate the correct state. In Chapter 2 reverse cascades can justify the Self-Confirming Equilibrium with unequal outcomes when discriminatory tastes have disappeared, since workers access aggregate information that might affect subjects' behavior over and above individual outcomes. Anderson and Holt (1997) provide evidence from the lab of the existence of such information cascades, and, to a minor extent, of reverse cascades. Hey and Morone (2002) show that lemmings survive even within a market contest.

The theoretical sunspot model postulates that agents believe that a variable, which is in fact unrelated to the economy, has real effects, and shows that such beliefs can induce the agents to behave in a manner that provides support for the postulated beliefs. Sunspots were introduced to the laboratory by Woodford (1990), who shows that cyclic sunspot equilibria can asymptotically emerge in an OLG framework when agents follow some adaptive learning schemes. Marimon, Spear and Sunder (1993) do not find evidence that sunspot equilibria exist when the extrinsic variable is not correlated with some real shock. However, they do find evidence that sunspots matter, taking the form of common past experience that influences agents behavior even when the real shock (correlated with sunspots) has been removed. This is a combination of hysteresis, i.e. the lagging of an effect behind its cause, and sunspots. What the experiment in section 3.3 tries to figure out is the existence of hysteresis without sunspots. In this case discriminatory tastes are the key variable that has real effects and that is withdrawn, while there is no extrinsic signal that drives the behavior of agents after the real shock disappears.

The paper by Bull, Schotter and Weigelt (1987), already mentioned when talking about asymmetric tournaments, is relevant also as far as the role of information is concerned. Giving contestants additional, but not complete, information about the actions of their opponents appears to slow the rate at which agents converge on their optimal choice. Section 3.4 shows that no significant differences are associated with the behavior of players accessing aggregate statistics in addition to individual outcomes.

3.6 Conclusions

This chapter is aimed at testing the predictions of a model that explores the role of workers' expectations of being discriminated against as an original explanation for the puzzling long-run persistence of observed discrimination against some minorities in the labor market. The model, presented in Chapter 2, provides a theoretical framework based on a two-stage game of incomplete information where preferences and beliefs of both sides of the labor market matter. In every constituent game two workers, one of whom is a minority worker, are drawn from their *ex ante* identical populations and randomly matched with one employer.[15] At the end of the first period the employer promotes one (and only one) worker after having observed the output they have produced, which is a function of their observable effort. Crucially, promotions depend via effort on workers' expectations about the unknown employer's type, which captures his possible disutility of promoting a minority worker. The importance of workers' expectations can be appreciated by

[15] What distinguishes the population of minority workers is an observable characteristic not related to their productivity (e.g. race, gender).

comparing the distribution of promotions across populations that arises when minority workers overestimate the percentage of employers characterized by tastes for discrimination with a situation in which such beliefs are correct *ceteris paribus.* This difference becomes crystal clear when there are actually only employers who do not discriminate against the minority either directly or statistically. Even in this circumstance unequal outcomes may emerge, caused by wrong beliefs of being discriminated against that are self-confirming. Minority groups who expect being discriminated against exert a lower effort on average, because of a lower expected return. This induces a lower observed percentage of promotions within minority workers, which in turn is consistent with their beliefs that there are employers characterized by discriminatory tastes.

The experiment replicates the model using a game where participants are randomly divided into two populations: red and blue. In every trial each participant has an endowment of 10 Euro cents and can decide how many cents to bid to get a prize worth 25 Euro cents. Bets are not given back to the players, neither to the winner nor to the loser, making the game equivalent to an all-pay auction. Therefore, at the end of the trial the winner gets 25 cents plus the amount not bet, while the loser only the amount not bet. In every lottery there are only two participants, one from the red population and one from the blue population. The players know that they face only one opponent in every trial and that it is possible to face the same opponent more than once during the same treatment, but of course they do not know when, given that random assignment takes place at the beginning of each trial. The prize is awarded to the higher bid and it is split if bids are equal, unless the opponents are assigned to a crazy computer which instead awards the prize to the red player regardless of the bids.

The mechanism underlying the SCE driven by wrong beliefs in the theoretical model is tested comparing the outcomes of an identical treatment proposed to the subjects both before and after they face an increasing fraction of crazy computers.

A reduction of bids made by disadvantaged subjects is observed in three out of seven sessions, in line with the theoretical predictions, but it vanishes rather quickly during the treatment, failing to generate the SCE driven by wrong beliefs about discrimination. The parallel of this finding in the labor market would be a situation in which minority workers, after having been discriminated against, expect that unfavorable conditions continue while biased employers have actually disappeared. Hence, minority workers reduce their effort, and accordingly they are promoted less frequently, but eventually they discover that biased employers have disappeared and balanced promotions across populations of workers are observed.

The main reason that the experiment does not provide strong evidence in favor of unequal outcomes driven by subjects' expectations, is that the experiment failed to separate the beliefs of the two populations. While in the theoretical model different expectations about the fraction of discriminatory employers are a necessary condition for unequal outcomes to emerge, in the lab Blues never expected a clearly higher fraction of crazy computers than did Reds. A possible explanation for this finding is that the experimental design laid too much emphasis on the discontinuity between treatments, preventing carryover effects from emerging.

Another possible reason that the experiment does not provide strong evidence in favor of unequal outcomes driven by subjects' expectations, is the fact that the set of choices is roughly continuous. This continuity means that there is only a very slight difference in the optimal behavior of the advantaged and the disadvantaged subjects, and therefore also a very low cost of experimenting to discover the true state of nature. Finally, it is also worth noting that, from a regression where all observable and unobservable individual characteristics as well as the parameters of the experiment are controlled for, Blues display a propensity to bid much more than Reds *ceteris paribus* in spite of the random assignment of the color.

The strategy method, on the other hand, supports the SCE driven by wrong beliefs. In fact, advantaged and disadvantaged subjects react in a different way when *ad hoc* aggregate statistics are displayed. A fictitious distribution shows a fraction of prizes won by the Blues decreasing from 80% to 0% in subsequent trials. Since all subjects are informed that there is no crazy computer, it might be expected that members of both populations ignore the aggregate statistics displayed and dissipate their endowment entirely. Red players, in fact, do not change their behavior significantly. Blue players, on the other hand, as the fictitious distribution of prizes becomes less and less favorable to them, lower their offers. As a result, the fraction of blue players bidding more than the average of their opponents also decreases from 74.3% in the first trial, when the distribution of prizes was supposed to be more favorable, to 25.7% in the last trial without any change in fundamentals. What happens in the last trial is that blue players are influenced by the aggregate statistics showing that none of them got the prize, and thus bid less. In turn, their lower bids make them less likely to win, leading to unequal outcomes that are consistent with wrong expectations that they were less likely to get the prize.

The experiment also deals with the relevance of the information structure by dividing the two populations in two subgroups. The first observes individual outcomes only (bids and winners of the games in which the player is directly involved), while the second is also informed about the distribution of bids within, as well as the distribution of promotion across, populations. From the experiment there is no evidence of different patterns between these two subgroups.

Concluding, findings of the experiment provide some evidence supporting the SCE driven by wrong beliefs, but the evidence cannot be considered fully satisfactory. In order to provide a more robust test of the theoretical model, future experiments should be modified to include such things as a dichotomized choice set or less clear-cut treatments. Another potentially fruitful way to test the model that is left for future research is to try a different version of the experiment in which Reds and Blues may behave differently even though they share similar beliefs about the fraction of crazy computers. Implementing a framework in which two Reds and two Blues compete for one prize could be a way to separate directly the behavior of Reds and Blues. In the version of the experiment presented in this chapter, similar beliefs lead to a similar behavior. If there are two players from each population, instead, even though all players share similar beliefs about the fraction of crazy computers, Reds should still bid a lot because, even though blue players are discriminated against, they need to compete against each other in order to get the prize, while for the Blues is pointless to try hard. Once behavior differs, a self-reinforcing mechanism could operate in which Blues bid less, they are not awarded the prize and they bid less and less, while Reds keep bidding significantly more.

Appendix: instructions

The Experiment - part A (shown at the beginning)

The experiment will last approximately 60 minutes, but the actual length depends on the speed of the slowest participant. The experiment is composed by two quizzes, eight stages and a questionnaire.

Numbers during the experiment represent Euro cents. Your final earnings will be the sum of all the Euro cents you earned throughout the experiment. Earnings depend on your choices as well as on the choices of your opponents during the game that will start in a few minutes.

The game consists of an auction, in which you have to bid in order to get a prize. The game will be repeated several times under potentially different conditions that will be explained at the beginning of each stage.

At the beginning of the experiment an algorithm will assign to every player a color label (red or blue) that will be effective for the whole experiment. The two populations (Red and Blue) will be of equal size. In every repetition of the game your opponent will be an anonymous player randomly drawn from the other population (i.e. if you belongs to the red population you will always play against a Blue and vice versa).

In every repetition of the game you will be endowed with 10 Euro cents. You have to decide how much to bid (from 0 to 10 cents) in order to win a prize worth 25 Euro cents.

- The higher bid wins the prize.
- If bids are equal, the prize is equally split.
- Bids are not given back, neither to the winner nor to the loser.

Your earnings in every repetition of the game depends on two factors:

1) The prize: 25 cents if you are awarded the prize, 12.5 cents if the prize is split, 0 if your opponent is awarded the prize;

2) How much of your endowment you did not bid.

N.B. You cannot save and transfer money from one repetition of the game to another. If you bid less than 10 cents the amount left will enter your earnings, but in the following repetition you will start again with 10 cents.

The Experiment - part B (shown after Treatment 1)

Now a different kind of computers is introduced into the game. These computers, which are called "crazy computers," award the prize always to the red player regardless of who bids more.

Notice that the computers we are talking about ("normal" vs. "crazy") are not the computers you have in front of you. The server computer has been programmed to receive the data from the client computers (i.e. to receive the bids that you enter in the PCs in front of you). In every repetition of the game, each pair of players is randomly associated with a partition of the server that can correspond to a normal computer (which assign the prize to the higher bid) or to a crazy computer (which always awards the prize to the Red).

NB: In every repetition of the game, the bids of each pair of players are randomly assigned to a partition of the server computer. Hence, being assigned to a crazy or to a normal computer during one repetition of the game, does not depend on the kind of computer faced in the previous repetition and does not predict anything about the kind of computer that will be faced in the following repetition. From this point of view it is like starting from the beginning at every repetition. You only have to bear in mind that each time there is a percentage of partitions of the server that represent crazy computers.

How many are the crazy computers? The percentage of crazy computers can vary from 0% to 100% during the experiment, but it is held constant within each stage of the experiment. In some stages you will know the fraction of crazy computers, in other stages that fraction will be unknown.

The introduction of crazy computers creates different conditions for Reds and Blues. To ensure that all participants have the same earnings opportunity, each Blue will receive a lump sum compensation of 3.5 Euro in addition to what earned during the experiment. This compensation corresponds to the estimated loss induced by crazy computers and DOES NOT depend on the choices you will make during the experiment.

Bibliography

[1] ANDERSON, D.M. AND M.J. HAUPERT (1999): "Employment and Statistical Discrimination: a Hands-On Experiment," *Journal of Economics*, 25:1, 85-102.

[2] ANDERSON, L.R. AND C.A. HOLT (1997): "Information Cascades in the Laboratory," *American Economic Review*, 87:5, 847-862.

[3] ANDERSON, L.R., R.G. FRYER AND C.A. HOLT (2002): "Experimental Studies of Discrimination," in *Handbook on Economics of Discrimination*, ed. by W. Rogers, forthcoming

[4] ANDERSON, S.P., J.K. GOEREE AND C.A. HOLT (1998): "Rent Seeking with Bounded Rationality: An Analysis of the All Pay Auction," *Journal of Political Economy*, 106:4, 828-853

[5] 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.

[6] BANERJEE, A.V. (1992): "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 107:3, 797-817.

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

[8] BICKCHANDANI, S., D. HIRSHLEIFER AND I. WELCH (1992): "A Theory of Fads, Fashion, Custom and Cultural Change as Informational Cascades," *Journal of Political Economy*, 100:5, 992-1026.

[9] BULL, C., A. SCHOTTER AND K. WEIGELT (1987): "Tournaments and Piece Rates: An Experimental Study," *Journal of Political Economy*, 95:1, 1-33.

[10] DAVIS, D.D. (1987). "Maximal Quality Selection and Discrimination in Employment," Journal of Economic Behavior and Organization, 8, 97-112.

[11] DAVIS, D.D. AND C.A. HOLT (1993): *Experimental Economics*. Princeton: Princeton University Press.

[12] DAVIS, D.D. AND R.J. REILLY (1994): "Do too Many Cooks Always Spoil the Stew? An Experimental Analysis of Rent-Seeking and the Role of a Strategic Buyer," Public Choice, 95, 89-115.

[13] DUFFY, J. AND J. OCHS (2003): "Cooperative Behavior and the Frequency of Social Interaction," mimeo, University of Pittsburgh.

[14] FERSHTMAN, C. AND U. GNEEZY (2001): "Some Experimental Evidence on the Evolution of Discrimination, Cooperation and Perceptions of Fairness," Economic Journal, 481, 679-703.

[15] FISCHBACHER, U. (2002): "zTree Experimenter's Manual," mimeo, University of Zurich.

[16] FRYER, R.G., J.K. GOEREE AND C.A. HOLT (2002a): "Experience-Based Discrimination," mimeo, The University of Virginia.

[17] ———————— (2002b): "An Experimental Test of Statistical Discrimination" mimeo, The University of Virginia.

[18] HARGREAVES-HEAP, S. AND Y. VAROUFAKIS (2002): "Some Experimental Evidence on the Evolution of Discrimination, Cooperation and Perceptions of Fairness," Economic Journal, 481, 679-703.

[19] HEY, J. AND A. MORONE (2002): "Do Markets Drive out Lemmings or Vice-Versa?," mimeo.

[20] LAZEAR, E.P AND S. ROSEN (1981): "Rank-Order Tournaments as Optimum Labor Contracts," Journal of Political Economy, 89:5, 841-64.

[21] LOHM, S. (2000) "Stereotypes, Sex Discrimination and Paranoia - an experimental study," Scandinavian WP in Economics 2000:1.

[22] MARIMON, R., S. SPEAR AND S. SUNDER (1993). "Expectationally-Driven Market Volatility: Experimental Evidence," Journal of Economic Theory, 61, 74-103.

[23] MURNINGHAM, J.K. (2001): "A Very Extreme Case of the Dollar Auction," mimeo, Kellogg Graduate School of Management, Northwestern University.

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

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

[26] ROTH, A.E. (1994). "Let's Keep the Con out of Experimental Econ.: A Methodological Note," *Empirical Economics*, 19, 99-109.

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

[28] SELTEN, R. (1967): "Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments," in *Beiträge zur Experimentellen Wirtschaftsforschung,* ed. by H. Sauermann. Tübingen: J.C.B. Mohr, pp. 136-168.

[29] SHUBIK, M. (1971): "The Dollar Auction Game: A Paradox in Non-Cooperative Behavior and Escalation," Journal of Conflict Resolution, 15109-111.

[30] WOODFORD, M. (1990). "Learning to Believe in Sunspots," Econometrica, 58:2, 277-308.

Chapter 4

Gender Wage Gap in Expectations and Realizations[1]

4.1 Introduction

This chapter focuses on the correspondence between wage expectations and wage realizations from a gender perspective using data collected among students and graduates of Bocconi University in Milan, Italy. We show that the gender gap implied by students' expectations one year after graduation is consistent with the gender gap implied by the earnings of their elder counterparts who have already graduated. There is instead a misperception of the gender gap ten years after graduation because students expect the gender gap to be roughly constant while realizations point toward an increasing gap with experience. The gender gap diminishes but does not disappear when several controls such as family background, place of birth, high school diploma, university program attended, performance at university, civil status and number of children are taken into account. Moreover, quite surprisingly, there is no evidence of a diminishing realized gender gap between subsequent cohorts of Bocconi graduates. On the contrary, the gender

[1]This chapter is part of a joint project with Andrea Ichino.

gap measured immediately after graduation shows a puzzling upward trend across cohorts.

When we distinguish between different levels of students' performance, the best students appear to be characterized by a significant gender gap at the beginning of their careers, which is underestimated in expectations but which remains approximately constant with experience. For the worse students, instead, the gender gap is smaller and correctly anticipated at the beginning of a career, but it increases significantly with their working life and this growth is not expected. These differences between the best and the worse students suggest that the careers of females are characterized by "glass ceilings" at high skill levels and by "sticky floors" at the opposite end of the skill spectrum.

The chapter is organized as follows. Section 4.2 describes the dataset. Section 4.3.1 presents evidence on the reliability of our data while Section 4.4 shows the econometric evidence and discusses the main results of this chapter in connection with the existing literature. Section 4.5 analyzes whether the perception of the gender wage gap differs according to students' performance. Concluding remarks follow in Section 4.6.

4.2 The dataset

Students' expectations were collected by circulating an anonymous questionnaire (reported at the end of the chapter) among second year Bocconi students. The questionnaire contains questions concerning wage expectations as well as personal information and family background. The data obtained were merged with Bocconi's administrative data about student's curricula. We obtained access to the same type of administrative information also for a sample of Bocconi graduates, who were interviewed about their current and past working situation. In this way a clean matching between similar Bocconi students and Bocconi graduates was made possible. Through this matching, we are now able to analyze the expectations of students comparing them with the realizations observed for similar graduates.

4.2.1 Students: expected working situation

The questionnaire concerning students' expectations was circulated by Bocconi staff attached to the yearly course evaluation forms. In this way most of the 2497 second year students received the questionnaire.[2] We got back complete reliable answers for 1154 questionnaires. Since the questionnaires were anonymous we had to use personal information like gender, date and province of birth to merge them with administrative data. Matching was successful for 887 observations. The remaining observations could not be merged either because of the incompleteness of the personal information (e.g. wage expectations, gender and/or date and/or province of birth missing) or because it was not possible to identify a unique counterpart of the questionnaire in the administrative data.

Despite this loss of observations, it is reassuring that descriptive statistics of the merged questionnaires do not significantly differ from those of the questionnaires which could not be merged. The only relevant difference that emerges is that among non-matched questionnaires there are fewer females (42.8% vs. 49.7%). Given that one of the purposes of this project is to check whether there are gender differentials in the expected working situation this might be a problem at first sight. However, breaking down matched and non-matched questionnaires across gender, the descriptive statistics of matched and non-matched males are similar. The same is true for the statistics of matched and non-matched females.[3] This evidence supports the fact that the missing information is missing almost randomly.

The questionnaire consists of three parts. The first part concerns the student's expected wage, occupation and sector of employment both one and ten years after

[2]Bocconi estimates that 75% of the students were attending the courses.

[3]These statistics are omitted to save space, but are available from the authors.

graduation. The second part focuses directly on the gender wage gap, asking students about the percentage gender wage differential they expect. In more detail, students are asked to use their expected wage (set equal to one hundred) as a benchmark and to report what is the wage that they think would be earned by a student with the same characteristics as theirs but of the other gender. In what follows, we will refer to this variable as the "self-reported" gender gap, to stress the difference with respect to the gender gap implied by the comparison between the average expected wages of males and females in the sample, collected in the first part of the questionnaire.

Moreover, students who give a "self-reported" gender gap different from zero are asked to choose among some possible explanations for such gap. The proposed explanations are closely linked with different theories in the discrimination literature (e.g. discriminatory tastes, statistical discrimination, human capital approach).

Finally, the role of the last section, which collects information about personal data and family background, is twofold. On the one hand, it makes it possible to merge the questionnaires with the administrative data. On the other hand, it provides a way of checking the reliability of the responses to the other parts of the questionnaire, because some answers, like those concerning school performance, can be verified using the administrative data.

4.2.2 Students and graduates: administrative data

As already mentioned, information taken from the questionnaires circulated among students was merged with Bocconi's administrative records. In addition to information about date of birth, place of birth, place of residence, etc. Bocconi's files keep track of students' high school background (name and place of the high school, type of diploma, grade obtained) and of all the details about students' university career (degree program; specialization; code, date and grade of all the passed exams). For the graduates, information about graduation (date, grade, etc.) is also available.

4.2.3 Graduates: working situation

A sample of Bocconi graduates was also interviewed, collecting information about their current and past working situation. This dataset contains a great deal of information, to be used also in other research projects. Here we use only the variables which are available also for students. It is important to keep in mind that the questionnaire circulated among Bocconi students has been designed to be compatible with the information available for graduates. Therefore, the questions have been designed to be as similar as possible to those asked in the survey of graduates. This latter was conducted by the Research Institute CIRM on behalf of Bocconi University. The target sample included all the 5091 graduates in four years: 1985, 1989, 1993 and 1997. CIRM selected a sample of 2802 students. However, 697 observations have been disregarded because they contain missing values for crucial variables like wages, leading to a final sample of 2105 observations.

4.3 Reliability of the data

4.3.1 Internal consistency of the expectations dataset

The dataset on student's expectations offers two ways of estimating the gender gap: first, using the average expected wages computed for the students of the two genders and, second, using the "self-reported" wage gap asked directly by the questionnaire. The comparison between these two different measures of the same concept allow us to check the internal consistency of the expectations dataset.

From descriptive statistics, which for reasons of space are not reported here, it emerges that the difference between the expected wage of males and females is greater than the "self-reported" gender gap. There are two possible explanations:

Table 4.1: Classes of income in the questionnaires

	up to 1032.91
from 1032.91	up to 2065.83
from 2065.83	up to 3098.74
from 3098.74	up to 4131.66
from 4131.66	up to 5164.57
from 5164.57	

Note: Monthly income net of taxes and contributions at 2001 prices (Euro).

either there is a misperception of the expected gender gap or the discrepancy is due to the fact that while the "self-reported" gender gap refers to identical students of different genders, in the sample males and females have different characteristics. However, if we include in a regression where the expected wage is the dependent variable the "self-reported" gender gap among the observable characteristics that are controlled for, the dummy for gender is no longer statistically significant. Similarly, this dummy is not significant in a regression with the same controls but where the sample has been restricted to students who show a "self-reported" gender gap equal to zero. We interpret these results as evidence in favor of the internal consistency of the dataset. In fact, controlling for several characteristics, differences between the expected wages across genders are not significantly different from the "self-reported" gender gap.[4]

4.3.2 Interval measures for wages

Both students and graduates were asked to report their expected and actual wage choosing between income classes according to the scheme described in Table 4.1.[5]

Regressions in this chapter are based on income measures obtained by assigning

[4]Results of this analysis are not reported to save space, but are available from the authors.

[5]The reason Table 4.1 involves Euro cents is that in the questionnaires amonunts were denominated in Lira, given that the questionnaires were circulated before the Euro became the official currency.

a point estimate to every class. To be precise we used the following rule. 75 percent of the upper bound was imputed for the lowest class; the mid point was imputed for each intermediate class and 125 percent of the lower bound was imputed for the highest class. This rule, like any other, is certainly arbitrary but it follows from the plausible assumption that the income distribution is uniform within each intermediate class, while being skewed toward the upper (lower) bound in the lowest (highest) class.

Robustness checks have been performed on this rule, with particular attention to the implications for gender differentials. Whenever possible, "Interval regressions" have been compared to OLS regressions based on income data constructed with this rule, obtaining very similar results. Note, however, that it was possible to compare these two type of results only when current wages were used.[6] Yet, the similarity of OLS and Interval regression results when both were feasible is reassuring.

4.4 Results

Table 4.2 shows the gender wage gap derived from:

1. students' wage expectations at one and ten years after graduation (respectively t_0 and t_1);

2. wages earned by the four cohorts of Bocconi graduates. For these cohorts t_0 stands for the first wage earned, while t_1 stands for the current (2001) wage. Note, therefore, that for graduates the number of years between t_0 and t_1 is

[6] In fact, the wage at the time of graduation needs to be corrected for the CPI because in the sample of graduates individuals started working in different years even within the same cohort. This causes the intervals to be different (and overlapping) across individuals. Furthermore, when wage growth is used interval regressions are either useless or not feasible regardless of the correction for the CPI.

Table 4.2: Percentage gender wage gap - uncontrolled

Cohort:	Students	grad85	grad89	grad93	grad97
N. obs:	887	234	469	637	765
t_0	-0.097***	-0.024	-0.030	-0.063**	-0.096***
	(0.027)	(0.053)	(0.032)	(0.029)	(0.026)
t_1	-0.139***	-0.569***	-0.423***	-0.263***	-0.158***
	(0.026)	(0.068)	(0.048)	(0.037)	(0.030)

Note: standard errors in parentheses with p<0.1 = *, p<0.05 = **, p<0.01 = ***.
The column labels "grad**" indicate the 19** cohort of graduates.
t_0= students' expected wage 1 year after graduation; 1st wage earned by graduates
t_1= students' expected wage 10 years after graduation; graduates' current wage
Percentage gap computed on real wages (prices in 2001).

not necessarily equal to nine. All the wages are real, having the wages in t_0 been corrected for the variation of the CPI.

Results in Table 4.2 are obtained, without controls, from the following basic regression

$$W_\tau^k = \alpha^k + \beta^k F + \varepsilon^k, \quad (4.1)$$

where W is the logaritm of expected or actual wages at $\tau = t_0$ or $\tau = t_1$, F is a dummy taking value one for females, ε is a disturbance term and k is the cohort. For each expected or actual wage measure the coefficient reported in the table is β, which approximates the percentage gender wage gap: for example -0.097 in the top left part of the table means that wage expectations of females are 9.7 percent lower than wage expectations of males. Note in this table, and in the similar following one, that figures in the t_0 row are comparable, being measures of the gender gap at the beginning of a career. Figures in the t_1 row are less easily comparable because they are measures of the gender gap at different levels of experience (10 years for students' expectations, 16 years for the 1985 cohort, etc.). However, comparisons within columns are possible as long as it is kept in mind that they give measures

Table 4.3: Percentage gender wage gap - controlled

Cohort:	Students	grad85	grad89	grad93	grad97
N. obs:	887	234	469	637	765
t_0	-0.082***	-0.009	-0.020	-0.064	-0.125***
	(0.027)	(0.121)	(0.064)	(0.044)	(0.029)
t_1	-0.106***	-0.308**	-0.150*	-0.185***	-0.162***
	(0.026)	(0.155)	(0.092)	(0.056)	(0.033)

Note: standard errors in parentheses with p<0.1 = *, p<0.05 = **, p<0.01 = ***.
Controls: high school diploma, family background, household business, degree program, place of birth, performance at university, civil status and number of children. For graduates, also part-time work is used as control.
The column labels "grad**" indicate the 19** cohort of graduates.
t_0= students' expected wage 1 year after graduation; 1st wage earned by graduates
t_1= students' expected wage 10 years after graduation; graduates' current wage
Percentage gap computed on real wages (prices in 2001).

of how the gender gap evolves during the working life in different cohorts and for different intervals of experience.

Four facts are immediately evident in Table 4.2. First, the gender gap expected by students one year after graduation is very similar to the gender gap experimented by the youngest cohort of graduates (1997). Second, the expected gender gap ten years after graduation seems to heavily underestimate the actual gender wage gap. The gap that students expect ten years after graduation is even lower than the gap that those who graduated in 1997 experience only four years after graduation. Third, the actual gender gap immediately after graduation shows a rather intriguing upward trend across subsequent cohorts: this gap is three times larger for the 1997 cohort than for the 1985 cohort. Fourth, both the actual and the expected gender gap increase with labor market experience, but the former seems to increase more.

How does the picture change when controls are included? Table 4.3 displays the percentage gender wage gap when several characteristics, like family background, place of birth, high school diploma, university program attended, performance at university, civil status and number of children are added to equation 4.1. The magnitude of the gender gap decreases in most of the cases when controls are included. This is not surprising given the important role played in particular by the civil status and the number of children in explaining the different achievement of males and females in the labor market. However, there is a remarkable exception: the youngest cohort. Although it is intuitive that the younger the cohort the smaller the importance of civil status and number of children as controls, it is striking that for those who graduated in 1997 the gender wage gap is even higher when controls are included. Moreover, the upward trend across cohorts in the gender gap immediately after graduation is still present when controls are included, which is a result worth particular attention.

As far as expectations are concerned, the expected gender gap one year after graduation is roughly correct even when individual characteristics are controlled for. Similarly, the inclusion of controls does not alter the finding that students heavily underestimate the expected gender gap ten years after graduation. Even when controls are included the gap that students expect ten years after graduation is lower than the gap that those who graduated in 1997 experience only four years after graduation.

As we said, the figures in the t_1 row are not directly comparable across columns because the time from first job is different for every cohort. Table 4.4 eliminates the problem by using the annual growth of expected and actual wages as a dependent variable. More specifically, the basic estimated equation is

$$\frac{\Delta W^k}{t_1 - t_0} = \delta^k + \gamma^k F + \eta^k, \qquad (4.2)$$

where, for every cohort k, ΔW^k is the difference of the logarithm of expected or actual wages between t_1 and t_0 and $t_1 - t_0$ is measured in years without rounding. For each expected or actual wage measure, the coefficient reported in the table is γ, which approximates the gender gap in the yearly growth of wages. The

Table 4.4: Annual growth of the gender wage gap

Cohort:	Students	grad85	grad89	grad93	grad97
N. obs:	887	234	469	637	765
uncontrolled	-0.017*	-0.081***	-0.079***	-0.064***	-0.049***
	(0.010)	(0.020)	(0.014)	(0.013)	(0.017)
controlled	-0.011	-0.054	-0.033	-0.038*	-0.036*
	(0.010)	(0.041)	(0.027)	(0.021)	(0.020)

Note: standard errors in parentheses with p<0.1 = *, p<0.05 = **, p<0.01 = ***
Controls: family background, place of birth, high school diploma, university program attended, performance at university, civil status and number of children. For graduates, also part-time work is used as control.
The column labels "grad**" indicate the 19** cohort of graduates.

first row of the table reports the uncontrolled estimates, while the second row reports results obtained by controlling for observable characteristics like family background, place of birth, high school diploma, university program attended, performance at university, civil status and number of children. For example -0.036 in the bottom right part of the table means that within the cohort of those who graduated in 1997 wages of females grew 3.6% less than wages of males for every year once controls are included. Table 4.4 confirms that students do not guess correctly the growth over time of the gender gap. In fact, they do not expect the gender gap to increase significantly, while this is an undisputable fact observable in the wage realizations of graduates.

Summarizing the findings of this section, a first important result is that in our data there is no evidence of a decreasing gender gap over time, i.e. across subsequent cohorts of Bocconi graduates, at the beginning of a career. On the contrary, our evidence points toward an increase of the gender gap in recent years, and in particular for the 1993 and 1997 cohorts.

This result is striking because several recent studies indicate that the gender wage gap has been narrowing since the '70s in most industrialized countries (see, for example, Blau and Kahn, 1996 and 1997). Specifically for the US, Datta Gupta, Oaxaca and Smith (2001) show a clear decline of the gender wage gap during the 1980s. As far as the United Kingdom is concerned, Blundell, Gosling, Ichimura and Meghir (2002) show that gender wage differentials for younger highly educated workers fell between 1978 and 1998. Similarly, Fitzenberger and Wunderlich (2002) find that the gender wage gap for full-time employed workers decreased considerably during the period 1975-95, particularly in the lower part of the wage distribution. Also in Italy the gender gap has been estimated to be narrowing by Flabbi (1997), who reports that gender differentials decreased from about 30% in 1977 to less than 20% in 1995. Only Scandinavian countries do not display a similar pattern, as reported by Datta Gupta, Oaxaca and Smith (2001), but in these countries the gender wage gap was already very small and is still among the lowest in the world.

Moreover, evidence of a narrowing gender gap between subsequent cohorts is somehow in line with the predictions of the most representative theoretical contributions within the discrimination literature. For example, both the discriminatory taste approach (Becker, 1957) and the statistical discrimination model (Arrow, 1973) imply that gender differentials should not survive in the long-run. A gender wage gap could persist in the long-run in the presence of self-confirming expectations, as suggested by Chapter 2. But even in this case there would be no reason to expect an increasing gap. Hence, it is rather puzzling to find such a pattern in the data analyzed in this chapter.

We can think of only one plausible reason explaining the difference of our results with respect to the literature. The increasing gender wage gap displayed in our data could be a consequence of the fading effects of the cost of living adjustment called *Scala Mobile* which prevailed in Italy during the '80s and which was abolished in 1992. As explained, for example, in Erickson and Ichino (1994), the design of this adjustment scheme implied a strong compressionary effect on wage differentials of all kinds and in particular on the gender wage gap. The abolition of the Scala Mobile is likely to have allowed an expansion of wage differentials which had been previously artificially compressed.

A second important result of this section is that the gender wage gap increases in the first part of the working life. In the literature, a few longitudinal studies provide evidence about the time profile of the gender gap within cohorts. Loprest (1992) finds an 11 percent gender wage gap at hiring within a sample of US young workers of all education levels during the period 1978-83. This gap increases in the first years after hiring and then decreases later during the working life. Light and Ureta (1995) present similar evidence. Kunze (2002) studies the evolution of the gender wage gap within the early stages of careers in Germany. She finds a gender gap of approximately 25 percent for the entry wages of skilled workers trained in vocational schools, but in contrast with the above studies, this gap remains roughly constant during the first eight years after hiring.

A third important result is that students' expectations appear to internalize correctly the existence of a gender gap at the beginning of the career, but not that such a gap increases during the working life. We are aware of only one paper to which this result can be compared, i.e. the paper by Brunello, Lucifora and Winter-Ebmer (2001), who collected a dataset containing information about wage expectations of more than 6000 European college students, although they do not have information on wage realizations and they are not interested in a gender perspective. The authors kindly gave us access to summary statistics of their data from which we could analyze wage expectations by gender. Focusing on the countries where at least 500 observations are available, the data suggest the existence of large differences in expectations across countries. The expected gender gap one year after graduation ranges from 9.2% in Switzerland to 18.7% in Germany, with Italy and Portugal situated in the middle with a gap of 12.6% and 16% respectively. As far as Italy is concerned, their figure differs from ours (9.7%

in Table 4.2), but it should be noticed that their sample of Italian students does not include Bocconi University. The cross country differences are less evident when looking at the expected gender gap ten years after graduation. In this case the gap is very similar in Germany and Switzerland (23.3% and 23.8%, respectively) while it is slightly higher in Portugal (26.5%) and Italy (28.3%). The corresponding figure in our dataset is 13.9%, which indicates that Bocconi students have expectations about the shape of the gender gap during the working life that significantly differ from the expectations of other Italian students.

4.5 Wage expectations and realizations at different levels of educational performance

In order to deepen our analysis of the relationship between wage expectations and wage realizations, in this section we stratify the sample of students and graduates according to their educational performance. Two slightly different stratification procedures have been used for graduates and students.

Using administrative data on the entire population of Bocconi graduates (i.e. not just the graduates interviewed by CIRM) the average grade that separates the top 25% and the bottom 25% of the population is used to define the best and the worst performing students. This is done separately for each cohort, since grades are likely to be comparable only within cohorts (see Table 4.5).

As far as students are concerned, again using population data from administrative records, the top and bottom 25% thresholds are identified according to a performance variable that summarizes the number of exams passed during the first year, weighted according to their difficulty and the grade obtained (see Table 4.6).

Table 4.7 shows the frequencies of top and bottom performers in our samples of graduates and students, using the thresholds defined above for the respective pop-

Table 4.5: Grades defining top and bottom 25% performance in the population of graduates

cohort	bottom 25%	top 25%
1985	< 24.37	> 27.15
1989	< 24.45	> 27.12
1993	< 25.20	> 27.82
1997	< 25.72	> 27.89

Note: the support of the grade variable is [18,30]

Table 4.6: Normalised indicator defining top and bottom 25% performance in the population of students

Students	bottom 25%	top 25%
2nd year	< 38.93	> 60.36

Note: the support of the performance variable is [0,100]

ulations. If the samples corresponded exactly to the populations these frequencies should always be equal to 25%, but this is clearly not the case. In particular, they are slightly smaller for the top group of graduates and this might be due to the fact that top graduates are more likely to earn higher wages and, for this reason, to be under-represented in the sample since they refused to answer the income question in the CIRM questionnaire. As far as students are concerned the frequency in the top group is almost 10 points higher than 25%, which may be due to the fact that students not attending classes are more likely to be worse performers and did not receive the questionnaire.[7]

On the basis of the stratification described above we have replicated the analysis of the previous section separately for top, intermediate and bottom students, in order to see whether the comparison between wage expectations and wage re-

[7]Remember that the students' questionnaires were attached to the evaluation forms of some courses. Therefore, our students sample has been drawn from the population of students still attending courses at the end of the term.

Table 4.7: Distribution of "top" and "bottom" in the sample

Group	% of "top"	% of "bottom"
Cohort 85	17.95	27.35
Cohort 89	21.11	24.31
Cohort 93	21.19	25.75
Cohort 97	23.01	24.71
2nd yr stud	34.22	12.86

Table 4.8: Expected and realized gender gap one year after graduation, by educational performance and controlling for observable characteristics

	top Stud.	top Grad.	medium Stud.	medium Grad.	bottom Stud.	bottom Grad.
N. obs:	288	452	466	1122	133	531
Female	-0.038	-0.165***	-0.087**	-0.119***	-0.134*	-0.138***
	(0.048)	(0.053)	(0.037)	(0.029)	(0.076)	(0.050)

Note: standard errors in parentheses with p<0.1=*, p<0.05=**, p<0.01=***.
Dependent variable: log of real wage.
Controls: family background, place of birth, high school diploma, university program attended, civil status and number of children. For graduates, also time since graduation and part-time work are used as control.

alizations differs according to educational performance. The analysis is performed pooling together the four cohorts of graduates because the sample size was not large enough to allow for the distinction between performance levels within each cohort separately. However, we include years since graduation into the regressions to control for experience.

Results obtained controlling for observable characteristics are presented in Table 4.8 for the gender gap measured immediately after graduation and in Table 4.9 for the yearly growth of this gap.

The first interesting fact emerging from these tables is that top students clearly underestimate the gender gap at the beginning of a career (t_0), while the guess of intermediate and bottom students is more accurate. This result hints at the

Table 4.9: Expected and realized gender gap growth, by educational performance and controlling for observable characteristics

	top Stud.	top Grad.	medium Stud.	medium Grad.	bottom Stud.	bottom Grad.
N. obs:	288	452	466	1122	133	531
Female	-0.008	-0.008	-0.002	-0.016*	0.005	-0.041***
	(0.005)	(0.016)	(0.004)	(0.009)	(0.008)	(0.015)

Note: standard errors in parentheses with p<0.1=*, p<0.05=**, p<0.01=***.
Dependent variable: real wage growth.
Controls: family background, place of birth, high school diploma, university program attended, civil status and number of children. For graduates, also time since graduation and part-time work are used as control.

possibility that top performing students may start their working career under the assumption that the human capital acquired in school is going to be the main determinant of success in the labor market, whereas the reality is different.

A second striking set of facts is offered by the comparison between graduates in the two tables. In the top performance group we see the largest gender gap at the beginning of the career (see Table 4.8). However, the growth of the gender gap during the working life is larger in the intermediate and bottom performance group (see Table 4.9).

These results suggest the possibility of different patterns of job assignment between males and females at different stages of a career and at different levels of the occupational hierarchy. For example, in the case of the top graduates, wage differences, possibly due to different job assignments, seem to emerge immediately after graduation and to persist more or less constantly during the career. A different pattern characterizes the bottom graduates. Here the evidence suggests that not only do wages differ already at the beginning of the career, but also that the difference increases during the working life. As a result, the gender wage gap increases with experience. Following Booth, Francesconi and Frank (1998), this evidence is consistent with the existence of "glass ceilings" for highly skilled female graduates, who are excluded from the very beginning of their careers from the same wage prospects offered to males of similar ability. At the same time, our evidence suggests that at the opposite end of the skill spectrum unskilled females experience "sticky floors" which prevent them from enjoying during their careers the same wage growth as their male counterparts.

4.6 Conclusions

The evidence presented in this chapter points towards some interesting findings. We show that the gender gap implied by students' expectations one year after graduation is consistent with the gender gap observed in the actual earnings of their older counterparts who have already graduated. There is instead a misperception of the gender gap ten years after graduation because students expect the gender gap to be roughly constant while realizations point toward an increasing gap with tenure. The gender gap diminishes but does not disappear when several controls such as family background, place of birth, high school diploma, university program attended, performance at university, civil status and number of children are taken into account.

A second set of intriguing results concerns the evidence on realized gender gaps independently of expectations. Here, in contrast with the recent literature for industrialized countries, we see no evidence of a diminishing gender gap between subsequent cohorts of Bocconi graduates at the beginning of a career. In particular, the gender gap immediately after graduation displays a puzzling upward trend and reaches particularly high and significant values in the most recent 1997 cohort. This result is likely to be a consequence of the elimination of the cost of living adjustment scheme called Scala Mobile which prevailed in Italy during the '80s and was abolished in 1992.

Finally, while the gender wage gap for the best graduates is large at the beginning of a career but remains more or less constant throughout the working life, for the worst graduates the gender gap starts slightly lower but increases more significantly with experience. These results suggest that the careers of females are characterized by "glass ceilings," in particular at high skill levels, and by "sticky floors" at the opposite end of the skill spectrum. Unfortunately, our data do not allow us to shed more light on the real nature and on the determinants of these differences in career developments.

The existence of gender differences of this kind, even in a very homogeneous group of highly skilled workers like that of by Bocconi graduates, is striking and clearly calls for more research and better data.

Bibliography

[1] 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.

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

[3] BLAU F.D., FERBER M.A. AND A.E. WINKLER (1998): *The Economics of Women, Men and Work*, Upper Saddle River, NJ: Prentice Hall.

[4] BLAU, F.D. AND L.M. KAHN (1996): "Wage Structure and Gender Earnings Differentials: an International Comparison," *Economica*, 63, S29-S62.

[5] BLAU, F.D. AND L.M. KAHN (1997): "Swimming Upstream: Trends in the Gender Wage differential in the 1980s," *Journal of Labor Economics*, 15, 1-42.

[6] BLUNDELL R., GOSLING A., ICHIMURA H. AND C. MEGHIR (2002): "Changes in the Distribution of Male and Female Wages Accounting for Employment Composition," mimeo.

[7] BOOTH A.L., M. FRANCESCONI AND J. FRANK (1998): "Glass Ceilings or Sticky Floors?," CEPR Discussion Paper 1965.

[8] BRUNELLO G., LUCIFORA C. AND R. WINTER-EBMER (2001): "The Wage Expectations of European College Students," mimeo.

[9] DATTA GUPTA N., OAXACA R.L. AND N. SMITH (2001): "Swimming Upstream, Floating Downstream: Trends in the U.S. and Danish Gender Wage Gap," CLS Working Papers, 01-6.

[10] ERICKSON C., AND A. ICHINO (1994): "Wage Differentials in Italy: Market Forces, Institutions and Inflation," in *Differences and Changes in the Wage Structure* ed. by Richard Freeman, and Larry Katz. Chicago, IL: NBER and University of Chicago Press.

[11] FITZENBERGER, B. AND G. WUNDERLICH (2002): "The Changing Gender Gap Across the Wage Distribution in the U.K.," mimeo.

[12] FLABBI L. (1997): "Gender discrimination and Return to Education: Estimations on Individual Data," *Rivista di Politica Economica*, 12.

[13] KUNZE A. (2002): "Gender Differences in Entry Wages and Early Career Wages," IZA Working Paper n. 626

[14] LIGHT A. AND M. URETA (1995): "Early-Career Work Experience and Gender Wage Differentials, Journal of Labor Economics," 13:1, 121-154.

[15] LOPREST P.J. (1992): "Gender Differences in Wage Growth and Job Mobility," *American Economic Review*, 82:2, 526-532.

Occupation and income expectations of Bocconi students

Dear Student, we kindly ask you to fill this questionnaire concerning entrance in the labor market. Please consider that data are collected for the sole purpose of scientific research and that results will be circulated referring to aggregate statistics only.

1. **After graduation do you expect to work in a household business?** ☐YES ☐NO

2. **Your occupation will more likely be:**

	1 year after graduation (one choice)	10 years after graduation (one choice)
Paid Employment		
White collar	☐	☐
Middle manager	☐	☐
General manager	☐	☐
Secondary school teacher	☐	☐
University teacher	☐	☐
Other paid employment (GIVE DETAILS)	____________	____________
Self Employment		
Business consultant	☐	☐
Professional (non business consultant)	☐	☐
Enterpreneur	☐	☐
Other self employment (GIVE DETAILS)	____________	____________

3. **In which sector? (one choice)**
☐manufacturing ☐finance ☐public ☐ trade ☐other (GIVE DETAILS)__________

4. **How much do you think your monthly labour income net of taxes and contributions will be (at constant prices)**

	1 year after graduation	10 years after graduation
Less than L. 2.000.000	☐	☐
L.2.001.000 - L.3.000.000	☐	☐
L.3.001.000 - L.4.000.000	☐	☐
L.4.001.000 - L.5.000.000	☐	☐
L.5.001.000 - L.6.000.000	☐	☐
L.6.001.000 - L.8.000.000	☐	☐
L.8.001.000 - L.10.000.000	☐	☐
More than L. 10.000.000	☐	☐

5. **Setting to 100 your wage 1 year after graduation, how much do you think would be earned by a student with the same characteristics as yours but of the other gender?**

☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
50	60	70	80	90	100	110	120	130	140	150

6. **Setting to 100 your wage 10 year after graduation, how much do you think would be earned by a student with the same characteristics as yours but of the other gender?**

☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
50	60	70	80	90	100	110	120	130	140	150

7. **If your answer in 5 and/or 6 was different from 100: why? (multiple choices allowed)**
☐ Characteristics and aptitudes actually differ between males and females
☐ Different distribution of household duties
☐ Employers expect differenct characteristics between males and females
☐ Employers' tastes given equal characteristics and hosehold duties

8. **Year** ☐ 2^ ☐ 3^ ☐4^ ☐ F.C.
9. **Degree Program** ☐CLE ☐CLEA ☐CLAPI ☐CLELI ☐CLG ☐CLEFIN ☐CLEACC
10. **Province of birth** ____________
11. **Gender** ☐ Male ☐ Female
12. **Date of birth** ________________
13. **Nr. of passed exams** ____
14. **Average grade** ______
15. **Education of the father** ☐PRIMARY ☐SECONDARY ☐COLLEGE ☐UNIVERSITY
16. **Education of the mother** ☐PRIMARY ☐SECONDARY ☐COLLEGE ☐UNIVERSITY
17. **Father's occupation** ________________________________
18. **Mother's occupation** ________________________________
19. **Tuition category** ☐ 1^ ☐ 2^ ☐3^ ☐ 4^ ☐5^ ☐6^

Thanks for your cooperation

[illegible]