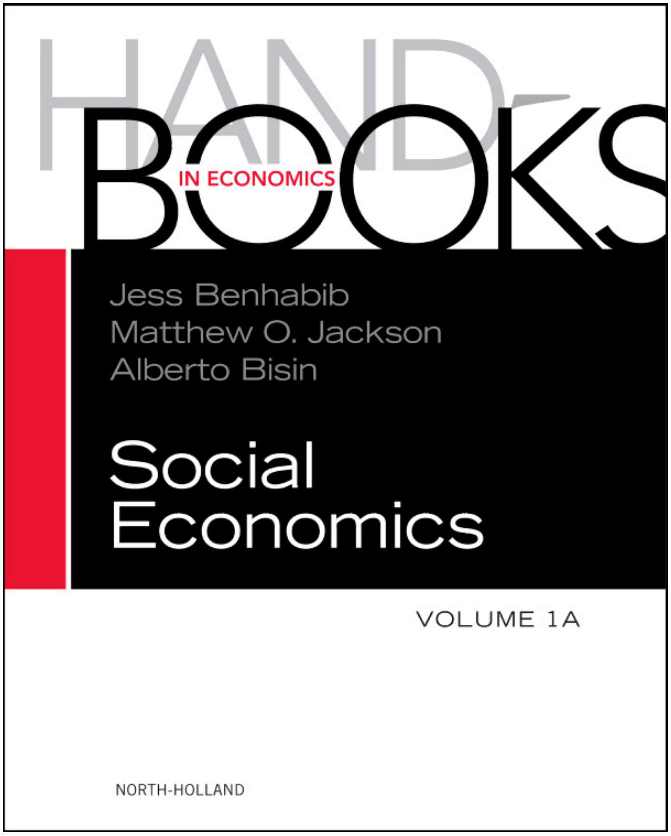

From: Hanming Fang and Andrea Moro, Theories of Statistical Discrimination and Affirmative Action: A Survey. In Jess Benhabib, Matthew O. Jackson and Alberto Bisin editors: Handbook of Social Economics, Vol. 1A,
The Netherlands: North-Holland, 2011, pp. 133-200.
ISBN: 978-0-444-53187-2

CHAPTER 5

Theories of Statistical Discrimination and Affirmative Action: A Survey*

Hanming Fang[§] and Andrea Moro[¶]

Contents

1. Introduction 134
2. The Use of Group Averages as a Proxy for Relevant Variables: the Exogenous Differences Literature 137
   2.1 A basic model of signal extraction 137
   2.2 Generating average group wage differentials 138
      2.2.1 Employers' risk aversion 138
      2.2.2 Human capital investment 139
      2.2.3 Tournaments 139
3. Discriminatory Outcomes as a Result of Coordination Failure 140
   3.1 Origin of equilibrium models of statistical discrimination 140
   3.2 Coate and Loury (1993a) 142
      3.2.1 The model 142
      3.2.2 Firms and workers' best responses 143
      3.2.3 Equilibrium 145
4. Discriminatory Outcomes Due to Inter-Group Interactions 147
   4.1 Discrimination as group specialization 147
      4.1.1 A model with production complementarities and competitive wages 147
      4.1.2 Cross-group effects 149
      4.1.3 The effect of group size 150
   4.2 Discrimination from search frictions 151
   4.3 Endogenous group formation 156
   4.4 Group interactions from peer effects 161
5. Dynamic Models of Discrimination 162
6. Affirmative Action 163
   6.1 A brief historical background 163
   6.2 Affirmative action in Coate and Loury (1993a) 165
      6.2.1 Modeling affirmative action 165
      6.2.2 Patronizing equilibrium: an example 166

* We thank Alberto Bisin, Jan Eeckhout, Yariv Fadlon, and Kevin Lang for providing comments on an initial draft of this chapter. We also thank our coauthors Glenn Loury and Peter Norman for providing many insights on statistical discrimination and affirmative action.

§ Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104; Duke University and the NBER. Email: hanming.fang@econ.upenn.edu

¶ Department of Economics, Vanderbilt University, 419 Calhoun Hall, Nashville, TN 37235. Email: andrea.moro@vanderbilt.edu

Handbook of Social Economics, Volume 1A
ISSN 0169-7218, DOI: 10.1016/S0169-7218(11)01005-7

6.3 General equilibrium consequences of affirmative action 170
6.4 Affirmative action in a two-sector general equilibrium model 173
6.5 Role model effects of affirmative action 176
6.6 Color sighted vs. Color blind affirmative action 178
6.6.1 Recent developments in the affirmative action policies related to college admission 178
6.6.2 Color sighted vs. Color blind affirmative action with exogenous skills 179
6.6.3 Color sighted vs. Colorblind affirmative action with endogenous skills 183
6.7 Additional issues related to affirmative action 189
7. Efficiency Implications of Statistical Discrimination 191
7.1 Efficiency in models with exogenous differences 191
7.2 Efficiency in models with endogenous differences 193
7.2.1 The planners' problem 193
7.2.2 The efficiency of discriminatory equilibria 195
8. Conclusion 195
References 197

Abstract

This chapter surveys the theoretical literature on statistical discrimination and affirmative action. This literature suggests different explanations for the existence and persistence of group inequality. This survey highlights such differences and describes in these contexts the effects of color-sighted and color-blind affirmative action policies, and the efficiency implications of discriminatory outcomes.

JEL Classification Codes: J150, J160, J700, J780

Keywords

Affirmative Action
Discrimination

1. INTRODUCTION

Statistical discrimination generally refers to the phenomenon of a decision-maker using observable characteristics of individuals as a proxy for unobservable, but outcome-relevant, characteristics. The decision-makers can be employers, college admission officers, health care providers, law enforcement officers, etc., depending on the specific situation. The observable characteristics are easily recognizable physical traits, which are used in the society to broadly categorize demographic groups by race, ethnicity, or gender. But, sometimes the group characteristics can also be endogenously chosen, such as club membership or language.

In contrast to taste-based theories of discrimination (see Becker 1957), statistical discrimination theories derive group inequality without assuming racial or gender animus, or preference bias, against members of a targeted group. In statistical discrimination models, the decision makers are standard utility or profit maximizers; and in most, though not all, models, they are also imperfectly informed about some relevant characteristics of the individuals, such as their productivity, qualifications, propensity to engage in criminal activity, etc., which rationally motivates the use of group statistics as proxies of these unobserved characteristics. While all models of statistical discrimination share these features, there exist important differences, which suggest different explanations for group inequality. This survey is structured to present these explanations and highlight these differences.[1]

The two seminal articles in this literature – Phelps (1972) and Arrow (1973) – which are often cited together, proposed in fact two different sources of group inequality. In Phelps (1972), and the literature that originated from it, the source of inequality is some unexplained exogenous difference between groups of workers, coupled with employers' imperfect information about workers' productivity. In the classic textbook example, if employers believe (correctly) that workers belonging to a minority group perform, on average, worse than dominant group workers do, then the employers' rational response is to treat differently workers from different groups that are otherwise identical. In another example, which is sometimes mentioned in labor economic textbooks, employers believe from past experience that young female workers have less labor market attachment than men, perhaps because of a higher propensity to engage in child-rearing. Therefore, they will be reluctant to invest in specific human capital formation of women, even if women are equally qualified as men. The employers' inability to observe individual's true labor market attachment forces them to rely on the group average. This makes it harder for women to achieve a higher labor market status. We survey this strand of the literature in Section 2.

In the literature that originated from Arrow (1973), average group differences in the aggregate are endogenously derived in equilibrium, without assuming any *ex-ante* exogenous differences between groups. Even in this strand of literature decision makers hold asymmetric beliefs about some relevant characteristic of members from different groups, but the asymmetry of beliefs is derived in equilibrium. This is why these beliefs are sometimes referred to as "self-fulfilling stereotypes". The typical approach in this literature is to design a base model with only one group that is capable of displaying multiple equilibria. When membership to "ex-ante" identical groups is added to the setup, between-group inequality can be sustained as an equilibrium outcome when the discriminated group fails to coordinate on the same equilibrium played by the dominant group. While there are always symmetric, "color-blind" equilibria in which groups behave identically, groups do not interact in these models. This feature, together with equilibrium multiplicity, makes coordination failure possible for one group. We describe these models in Section 3.

[1] For earlier surveys of the related literature with a stronger emphasis on empirical research, see Cain (1986) and Altonji and Blank (1999).

Coordination failure is not the only source of inequality in models with self-fulfilling stereotypes. A recent strand of literature, which we describe in Section 4, emphasizes inter-group interactions in models with complementarities (for example in production technology). Asymmetric equilibria are possible where *ex-ante* identical groups specialize in tasks that have different marginal productivity. These equilibria may exist even when there is a unique symmetric equilibrium. Because of the complementarities, in this class of models there are conflicting interests among groups regarding issues such as affirmative action. Section 4 will also present a model where group inequality emerges as a result of job search frictions instead of informational frictions, and a model where group identities, as well as skill investment decisions, are endogenously chosen.

Most of these models, with some exceptions, are not designed to explain which group ends up being discriminated. Groups are *ex-ante* identical; therefore the focus of these theories lies more in trying to explain the persistence of inequality, rather than its origins, which are implicitly assumed to be based on historical factors. These considerations are more appropriately studied by dynamic models. We survey the small dynamic statistical discrimination literature in Section 5.

In Section 6, we will look at different policy implications from these models, in particular using the models with self-fulfilling stereotypes. Outcome-based policies, such as affirmative action quotas, or the application of disparate impact tests, seem particularly suited to eliminate inequality based on self-fulfilling stereotypes. If the imposition of the quota can eliminate the asymmetric discriminatory equilibria and lead different groups to coordinate on a symmetric outcome, then the temporary affirmative action policy might eliminate inequality. Typically, however, the literature finds that outcomes where inequality persists will remain possible, despite the fulfillment of the policy requirements. While policies may be designed so that only symmetric outcomes remain after their applications, such policies are typically dependent on special modeling assumptions. We also review in this section some interesting theoretical analysis that compares the "color-sighted" and "color-blind" affirmative action policies in college admissions.

Finally, Section 7 presents some considerations regarding the efficiency properties of discriminatory outcomes in statistical discrimination models, and Section 8 concludes.

The concept of statistical discrimination has been applied mostly to labor market examples where employers discriminate against one group of workers. This is why this survey presents mostly labor market related examples, but the reader is advised to consider that the same concepts and theories are applicable to other markets and socio-economic situations. We have chosen for convenience to use racial discrimination of *W(hites)* against *B(lacks)* as the running example because this has been the choice in most of the literature. This choice of notation should not be interpreted as implying that other examples are less relevant, or that racial inequality is the most relevant application of all the theories this survey will describe.

2. THE USE OF GROUP AVERAGES AS A PROXY FOR RELEVANT VARIABLES: THE EXOGENOUS DIFFERENCES LITERATURE

In this section, we describe a simple model where group identity serves as a proxy for unobserved variables that are relevant to economic outcomes. We begin with describing a version of the seminal model of statistical discrimination by Phelps (1972). This model generates inequality from different sources, depending on the details of how the labor market is modeled, and on the nature of the groups' intrinsic differences.

2.1 A basic model of signal extraction

Consider the example of an employer that does not observe with certainty the skill level of her prospective employees, but observes group identity $j \in \{B, W\}$. Workers' skill q is assumed to be equal to the value of their marginal product when employed, and is drawn from a normal skill distribution $N(\mu_j, \sigma_j^2)$. Employers observe group identity and a noisy signal of productivity, $\theta = q + \varepsilon$, where ε is a zero-mean error that is normally distributed according to $N(0, \sigma_{\varepsilon j}^2)$.

In a competitive labor market where all employers share the same type of information, workers are paid the expected productivity conditional on the value of the signal. Each employer infers the expected value of q from θ using the available information, including group identity. The skill and the signal are jointly normally distributed, and the conditional distribution of q given θ is normal with mean equal to a weighted average of the signal and the unconditional group mean (see DeGroot 2004):

$$E(q|\theta) = \frac{\sigma_j^2}{\sigma_j^2 + \sigma_{\varepsilon j}^2}\theta + \frac{\sigma_{\varepsilon j}^2}{\sigma_j^2 + \sigma_{\varepsilon j}^2}\mu_j \tag{1}$$

Intuitively, if the signal is very noisy (that is, if the variance of ε is very high), the expected conditional value of workers' productivity is close to the population average regardless of the signal's value. At the other extreme, if the signal is very precise ($\sigma_{\varepsilon j}$ is close to zero), then the signal provides a precise estimate of the worker's ability.

Phelps (1972) suggested two cases that generate inequality, which is implicitly defined as an outcome where two individuals with the same signal, but from different groups, are treated differently.

Case 1. In the first case, assume that groups' signals are equally informative, but one group has lower average human capital investment, that is, $\sigma_{\varepsilon B} = \sigma_{\varepsilon W} = \sigma_\varepsilon$, and $\sigma_B = \sigma_W = \sigma$, but $\mu_B < \mu_W$. In this case, B workers receive lower wages than W workers with the same signal, because employers rationally attribute them lower expected productivity, after observing they belong to a group with lower productivity.

Case 2. In the second case, the unconditional distributions of skills are the same between the two groups ($\sigma_B = \sigma_W = \sigma$, and $\mu_B = \mu_W = \mu$), but the signals employers receive are differently informative, e.g., $\sigma_{\varepsilon B} > \sigma_{\varepsilon W}$.[2] From this assumption, it follows that B workers with high signals receive lower wages than same-signal workers from the W group, and the opposite happens to workers with low signals.

While this basic model is capable of explaining differential treatment for same-signal workers from different groups, on average workers of the two groups receive the same average wage, unless average productivity is assumed to be exogenously different as in Case 1, which is not an interesting case from a theoretical perspective.

Note also that in this model all workers are paid their expected productivity conditional on available information. Thus, differential treatment of same-signal workers from different groups does not represent "economic discrimination," which is said to occur if two workers with identical (expected) productivity are paid differently.[3,4]

2.2 Generating average group wage differentials

In this section, we present various extensions of Phelps' model that generate different group outcomes. All of these extensions are based on Phelps' "Case 2" assumption of different signal informativeness across groups.[5]

2.2.1 Employers' risk aversion

Aigner and Cain (1977) proposed to incorporate employers' risk aversion into the standard Phelps' setup. Assuming, for example, that employers' preferences are given by:

$$U(q) = a + b\exp(-cq),$$

then employers' expected utility from hiring a worker with signal θ is given by:

$$E(U(q)|\theta) = a - b\exp\left[-cE(q|\theta) + \frac{c}{2}Var(q|\theta)\right].$$

From the properties of the conditional normal distribution we have:

$$Var(q|\theta) = \frac{\sigma_j^2\sigma_{\varepsilon j}^2}{\sigma_j^2 + \sigma_{\varepsilon j}^2},$$

which is increasing in $\sigma_{\varepsilon j}$. This implies that wages are decreasing in $\sigma_{\varepsilon j}$. Therefore the group with the higher noise (e.g., B workers if $\sigma_{\varepsilon B} > \sigma_{\varepsilon W}$) receives, on average, a lower wage. Employers are compensated for the risk factor incorporated in each B worker's higher uncertainty in productivity, measured by the term $cVar(q\,|\,\theta)/2$.

[2] This assumption can be rationalized assuming some communication of language barriers between employers and minorities, see, Lang (1986).

[3] See Stiglitz (1973) and Cain (1986) for early distinctions between statistical and economic discrimination.

[4] In Mailath, Samuelson and Shaked (2000) discussed in Section 4.2, differential treatment of workers with different races features economic discrimination.

[5] An example of an extension to "Case 1" is Sattinger (1998), where it is assumed that groups are homogenous in productivity but their workers differ in the probabilities of quitting their jobs. Firms observe quit rates imperfectly and profit maximization leads them to set unequal employment criteria or unequal interview rates across groups.

2.2.2 Human capital investment

Lundberg and Startz (1983) adopted a different approach, which was later exploited by the literature we will review in Sections 3 and 4. They assumed that worker's productivity q is partly determined by a costly human capital investment choice the worker undertakes before entering the labor market. Specifically, they parameterize $q = a + bX$, where X is human capital investment, b is a parameter common to all workers, and a is drawn from a normal distribution with mean μ and variance σ^2, common to groups B and W. The investment cost is a convex function $C(X) = cX^2/2$. After the human capital investment decision is made, the labor market works as in Case 2 of Phelps' model, that is, groups are assumed to differ in the information of the signal of productivity. Specifically, workers from group j with productivity q receive a signal $\theta = q + \varepsilon_j$ where as before ε_j is drawn from a Normal density $N(0, \sigma^2_{\varepsilon j})$.

Following (1), group j workers choose human capital investment to solve:

$$\max_{X_j} \int E(q|\theta)d\theta - C(X_j)$$
$$= \max_{X_j} \int \frac{\sigma^2}{\sigma^2 + \sigma^2_{\varepsilon j}}(a + bX_j + \varepsilon_j)d\varepsilon_j + \frac{\sigma^2_{\varepsilon j}}{\sigma^2 + \sigma^2_{\varepsilon j}}\mu - \frac{1}{2}cX_j^2.$$

Thus group j workers' optimal human capital investment is:

$$X_j^* = \frac{b}{c}\frac{\sigma^2}{\sigma^2 + \sigma^2_{\varepsilon j}}, \tag{2}$$

that is, members of the group with the higher signal noise invest less than members from the group with the lower signal noise.[6] Assuming for example that $\sigma^2_{\varepsilon B} > \sigma^2_{\varepsilon W}$, then in the labor market outcome workers from group B receive lower wages, on average, than workers from group W despite sharing the same distribution of *ex- ante* human capital endowment a. This outcome clearly relies on the existence of some form of heterogeneity across groups, namely, the signal informativeness.

2.2.3 Tournaments

Cornell and Welch (1996) embedded Phelps' "Case 2" assumption in a tournament model. Their observation was that if one group has a more informative signal, then this group's variance of the expected productivity is higher. For example, using Phelps' simple parameterization, workers with signal greater than the average have higher expected productivity if the signal is more precise, whereas the opposite is true for workers with a signal lower than their expected productivity. If labor demand is limited

[6] A version of this model can be written with heterogeneous investment costs. Moro and Norman (2003b) use this parameterization to generate log-normally distributed wages in equilibrium, which are suitable for empirical investigation.

compared to supply (e.g., the pool of candidates for a job is larger than the number of positions available), then jobs will go to the candidates with higher signals. Even if groups receive the same signals on average, the probability that the best signals belong to candidates from the dominant group is higher, which generates group inequality.

This intuition carries to more general paremeterizations. Cornell and Welch (1996) model information by assuming that many signals of productivity are available, all drawn from the same distribution, and assume that members of the dominant group can send employers a larger number of signals than members of the discriminated group. They prove that for any underlying signal distribution, the variance of the expected productivity is higher for the dominant group. As the number of candidates relative to the number of spots increase, the probability that members of the dominant group fill all positions approaches one.

3. DISCRIMINATORY OUTCOMES AS A RESULT OF COORDINATION FAILURE

In the models reviewed in Section 2, race, gender, or any group affiliation, is used in the determination of wages by firms in the competitive market because the distribution of signals about workers' productivity exogenously depends on the group identities. In this section, we review the literature that derives group differences endogenously even when groups share identical fundamentals. Outcomes with inequality can be thought of as the result of a self-fulfilling prophecy, and can be interpreted as group-wide coordination into the different equilibria of a base model in which group identity is ignored.

3.1 Origin of equilibrium models of statistical discrimination

Arrow's (1973) paper laid out the ingredients for a theory of discriminatory outcomes based on "self-fulfilling prophecies" with endogenous skill acquisition. First, the employers should be able to freely observe a worker's race. Second, the employers must incur some cost before they can determine the employee's true productivity (otherwise, there is no need for the use of surrogate information such as race or gender). Third, the employers must have some preconception of the distribution of productivity within each of the two groups of workers.

Arrow proposed the following model. Suppose that each firm has two kinds of jobs, skilled and unskilled, and the firms have a production function $f(L_s, L_u)$ where L_s is skilled labor and L_u is the unskilled labor. Denote with f_1 and f_2 the first derivatives of f with respect to the first and second arguments, respectively. All workers are qualified to perform the unskilled job, but only skilled workers can perform the skilled job.

Skills are acquired through investment. Workers have skill investment cost c, which is distributed in the population according to the cumulative distribution function $G(\cdot)$ which does not depend on group identity. Suppose that a proportion π_W of whites and a proportion of π_B of blacks are skilled, which will be determined in equilibrium. In order to endogenize the skill investment decisions, Arrow proposed the following model of wage differences between the skilled and unskilled jobs. Suppose that workers are assigned either to the skilled job or to the unskilled job. If a worker is assigned to the unskilled job, she receives a wage $w_u = f_2\ (L_s,\ L_u)$, independent of the race group of the worker. If a worker is assigned to the skilled job, then Arrow assumes that the worker will receive a wage contract that pays a group $j \in \{B,\ W\}$ worker wage $w_j > 0$ if that worker is tested to be skilled and 0 otherwise. Finally, the firm must pay a cost r to find out whether or not the worker is skilled. Arrow claims that competition among firms will result in a zero profit condition, therefore,

$$r = \pi_W[f_1(L_s, L_u) - w_W],$$
$$r = \pi_B[f_1(L_s, L_u) - w_B].$$

These imply that:

$$w_W = \frac{\pi_B}{\pi_W} w_B + \left(1 - \frac{\pi_B}{\pi_W}\right) f_1(L_s, L_u).$$

Note that if for some reason $\pi_B < \pi_W$, then $w_B < w_W$. Thus, blacks will be paid a lower wage in the skilled job if they are believed to be qualified with a lower probability. As a result, Arrow (1973) shifted the explanation of discriminatory behavior from preferences to beliefs.

Arrow then provided an explanation for why π_W and π_B might differ in equilibrium even though there are no intrinsic differences between groups in the distribution of skill investment cost $G\ (\cdot)$. Workers invest in skills if the gains of doing so outweigh the costs. Arrow takes the gains to be $w_j - w_u$ for group j workers.[7] Given the distribution of skill investment cost $G\ (\cdot)$, the proportion of skilled workers is $G\ (w_j - w_u)$, namely the fraction of workers whose skill investment cost c is lower than the wage gain from skill investment $w_j - w_u$. Equilibrium requires that:

$$\pi_j = G(w_j(\pi_W, \pi_B) - w_u), \text{ for } j \in \{B, W\}. \tag{3}$$

In a symmetric equilibrium, $\pi_W = \pi_B$, and in an asymmetric equilibrium, $\pi_B \neq \pi_W$. Arrow then notes that the system (3) can have symmetric as well as asymmetric equilibria. The intuition for the asymmetric equilibria is simple: if very few workers invest in a particular group, the firms will rationally perceive this group as unsuitable for the skilled task and equilibrium wages for this group in the skilled job will be low, which will in turn give little incentive for the workers from this group to invest. That is, self-fulfilling prophecies can lead to multiple equilibria. If groups coordinate on different

[7] Note that this is not entirely consistent with the labor market equilibrium conditions. Because $w_u > 0$, and any unqualified worker who is hired on the skilled job will eventually get a wage 0, no unqualified worker should agree to be hired on the skilled job in the first place.

equilibria, then discrimination arises with one group acquiring less human capital and receiving lower wages than the other group.[8]

3.2 Coate and Loury (1993a)

Coate and Loury (1993a) presented an equilibrium model of statistical discrimination where two *ex ante* identical groups may end up in different, Pareto ranked, equilibria. Coate and Loury's model formalizes many of ideas that were originally presented loosely in Arrow (1973), but it assumes that wages are set exogenously from the model.[9] The key element of Coate and Loury's model is that a worker's costly skill investment may not be perfectly observed by firms. Thus, firms may rely on the race of the worker as a useful source of information regarding the worker's skill. This introduces the possibility of self-fulfilling equilibria. If the firms believe that workers from a certain racial group are less likely to be skilled, and thus impose a higher threshold in assigning these workers to higher paying jobs, it will indeed be self-fulfilling to lower these workers' investment incentives, which in turn rationalizes the firms' initial pessimistic belief. Analogously, more optimistic belief about a group can be sustained as equilibrium. This is the source of multiple equilibria in Coate and Loury model. Discriminatory outcomes arise if two groups of identical workers play different equilibria.

As in Arrow's model, *ex ante* discrimination is generated by "coordination failure." It is important to emphasize that in this model there are no inter-group interactions, other than possibly when affirmative action policies such as employment quotas are imposed (see Section 6). In contrast, in the models we discuss in Section 4, inter-group interaction is the key mechanism for discriminatory outcomes for *ex ante* identical groups.

3.2.1 The model

Consider an environment with two or more competitive firms and a continuum of workers with unit mass. The workers belong to one of two identifiable groups, B or W, with $\lambda \in (0, 1)$ being the fraction of W in the population.

Firms assign each worker into one of two task that we respectively label as "complex" and "simple". Coate and Loury assume that wages on the two tasks are exogenous and are as follows: a worker receives a net wage ω if he is assigned to the complex task, and 0 if he is assigned to the simple task. The firm's net return from workers, however, depends on the workers' qualifications and their assigned task, which are summarized in Table 1. Thus the qualification is important for the complex task, but not for the simple task.

Workers are born to be unqualified, but they can become qualified if they undertake some costly *ex-ante* skill investment. Suppose that the cost of skill investment,

[8] Spence (1974) also suggested an explanation for group inequality based on multiple equilibria in his classic signaling model.

[9] This assumption can be relaxed in a model of linear production technology without affecting any of the main insights. New economic insights emerge if wages are endogenized in a model with nonlinear production technology. See Moro and Norman (2003a, 2004) described in Section 4.1.

Table 1 Firms' net return from qualified and unqualified workers in the complex and simple tasks

Worker\Task	**Complex**	**Simple**
Qualified	$x_q > 0$	0
Unqualified	$-x_u < 0$	0

denoted by c, is heterogenous across workers and is distributed according to cumulative distribution function (CDF) $G(\cdot)$, which is assumed to be continuous and differentiable. Importantly, $G(\cdot)$ is group independent: workers from different groups share the same cost distribution.

The most crucial assumption of the model is that workers' skill investment decisions are unobservable by the firms. Instead, firms observe a noisy signal $\theta \in [0, 1]$ of the worker's qualification. We assume that the signal θ is drawn from the interval [0, 1] according to PDF $f_q(\theta)$ if the worker is qualified, and according to $f_u(\theta)$ if he is unqualified. The corresponding CDF of f_q and f_u are denoted by F_q and F_u, respectively. To capture the idea that the noisy signal θ is informative about the workers' qualification, we assume that the distributions $f_q(\cdot)$ and $f_u(\cdot)$ satisfy the following Monotone Likelihood Ratio Property (MLRP):

Assumption 1. ***(MLRP)*** *$l(\theta) \equiv f_q(\theta) / f_u(\theta)$ is strictly increasing and continuous in θ for all $\theta \in [0, 1]$.*

It is useful to observe that this assumption is without loss of generality: for any pair of distributions f_q and f_u, we can always rank the signals according to the ratio $f_q(\theta)/f_u(\theta)$ and re-label the signals in accordance to their rankings. As we will see below, the MLRP assumption has two important and related implications. First, it implies that qualified workers, i.e., workers who have invested in skills, are more likely than unqualified workers to receive higher signals; second, it also implies that the posterior probability that a worker is qualified is increasing in θ.

The timing of the game is as follows. In Stage 1, Nature draws workers' types, namely, their skill investment cost c from the distribution $G(\cdot)$; in Stage 2, workers, after observing their type c, make the skill investment decisions, which are not perfectly observed by the firms; instead, the firms observe a common test result $\theta \in [0, 1]$ for each worker drawn respectively from PDF $f_q(\cdot)$ or $f_u(\cdot)$ depending on the worker's skill investment decision; finally, in Stage 3, firms decide how to assign the workers to the complex and simple tasks.

3.2.2 Firms and workers' best responses

The equilibrium of the model can be solved from the last stage. To this end, consider first the firms' task-assignment decision. Suppose that a firm sees a worker with signal θ from a group where a fraction π has invested in skills. The posterior probability that such a worker is qualified, denoted by $p(\theta; \pi)$, follows from Bayes' rule:

$$p(\theta;\pi) = \frac{\pi f_q(\theta)}{\pi f_q(\theta) + (1-\pi) f_u(\theta)}. \tag{4}$$

This updating formula, (4), illustrates a crucial insight: in environments with informational frictions (because workers' skill investment decisions are not perfectly observed by the firms), firms' assessment about the qualification of *a particular worker* with test signal θ depends on their prior about the fraction of *the group* that has invested in skills, i.e., π. Hence, a worker's investment not only increases her own chances of obtaining higher signals and higher expected wages, but also increases the employers' prior of all workers from the same group. This informational externality is the key source of the multiplicity of equilibria in this model.

Now consider the firm's task assignment decision in Stage 3 of a worker with a test signal θ belonging to a group where a fraction π have invested in skills. Using Table 1, the firm's expected profit from assigning such a worker to the complex task is:

$$p(\theta;\pi)x_q - [1 - p(\theta;\pi)]x_u, \tag{5}$$

because with probability $p(\theta;\pi)$ the worker is qualified and will generate x_q for the firm, but with probability $1 - p(\theta;\pi)$ he is unqualified and will lead to a loss of x_u if he is mistakenly assigned to the complex task. On the other hand, if such a worker is assigned to the simple task, the firm's profit is 0. Thus, the firm will optimally choose to assign such a worker to the complex task in Stage 3 if and only if:

$$p(\theta;\pi)x_q - [1 - p(\theta;\pi)]x_u \geq 0. \tag{6}$$

Using the expression (4) for $p(\theta;\pi)$, (6) is true if and only if:

$$\frac{f_q(\theta)}{f_u(\theta)} \geq \frac{1-\pi}{\pi}\frac{x_u}{x_q}. \tag{7}$$

Because of the MLRP assumption that f_q/f_u is monotonically increasing in θ, (7) holds if and only if $\theta \geq \widetilde{\theta}(\pi)$ where the threshold $\widetilde{\theta}(\pi)$ is determined as follows. If the equation:

$$\frac{f_q(\theta)}{f_u(\theta)} = \frac{1-\pi}{\pi}\frac{x_u}{x_q} \tag{8}$$

has a solution in (0,1), then $\widetilde{\theta}(\pi)$ is the unique solution (where the uniqueness follows from the MLRP); otherwise, $\widetilde{\theta}(\pi) = 0$ if $f_q(0)/f_u(0) \geq (1-\pi)x_u/(\pi x_q)$, and $\widetilde{\theta}(\pi) = 1$ if $f_q(1)/f_u(1) \leq (1-\pi)x_u/(\pi x_q)$. It is also clear that whenever the threshold $\widetilde{\theta}(\pi) \in (0,1)$, we have

$$\frac{d\widetilde{\theta}}{d\pi} = -l'(\widetilde{\theta}(\pi))\frac{x_u}{x_q}\frac{1}{\pi^2} < 0, \tag{9}$$

where $l(\theta) \equiv f_q(\theta)/f_u(\theta)$. That is, as the prior probability that a worker is qualified gets higher, the firms use a lower threshold of the signal in order to assign a worker to the complex task.

Now we analyze the workers' optimal skill investment decision at Stage 2, given the firms' sequentially rational behavior in Stage 3 as described above.

Suppose that in Stage 3, the firms choose a task assignment that follows a cutoff rule at $\tilde{\theta}$. If a worker with cost c decides to invest in skills, he expects to be assigned to the complex task, which pays $\omega > 0$, with probability $1 - F_q(\tilde{\theta})$ which is the probability that a qualified worker will receive a signal above $\tilde{\theta}$ (recall that F_q is the CDF of f_q). Thus his expected payoff from investing in skills in Stage 2 is:

$$\left[1 - F_q(\tilde{\theta})\right]\omega - c. \tag{10}$$

If he does not invest in skills, the signal he receives will nonetheless exceed $\tilde{\theta}$, and thus will be mistakenly assigned to the complex task with probability $1 - F_u(\tilde{\theta})$ (recall that F_u is the CDF of f_u). Hence his expected payoff from not investing in skills is:

$$\left[1 - F_u(\tilde{\theta})\right]\omega. \tag{11}$$

Hence, a worker with cost c will invest if and only if:

$$c \leq I(\tilde{\theta}) \equiv \left[F_u(\tilde{\theta}) - F_q(\tilde{\theta})\right]\omega. \tag{12}$$

The term $I(\tilde{\theta}) \equiv \left[F_u(\tilde{\theta}) - F_q(\tilde{\theta})\right]\omega$ denotes the benefit, or incentive, of the worker's skill investment as a function of the firms' signal threshold $\tilde{\theta}$ in the task assignment decision. A few observations about the benefit function $I(\cdot)$ can be useful. Note that:

$$I'(\tilde{\theta}) = \omega\left[f_u(\tilde{\theta}) - f_q(\tilde{\theta})\right] > 0 \tag{13}$$

if, and only if $l(\tilde{\theta}) < 1$. Because $l(\cdot)$ is assumed to be monotonic, it immediately follows that $I(\cdot)$ is a single peaked function. Moreover, $I(0) = I(1) = 0$. That is, if the firm assigns all signals (the case $\tilde{\theta} = 0$), or if the firm assigns no signals (the case $\tilde{\theta} = 1$) to the complex task, then workers will have no incentive to invest in skills. Figure 1 depicts one possible function $I(\cdot)$ satisfying these properties.

3.2.3 Equilibrium

Given the workers' optimal investment rule in response to the firms' assignment threshold $\tilde{\theta}$ as specified by (12), the fraction of workers who rationally invests in skills *given a cutoff* $\tilde{\theta}$ is simply the measure of workers whose investment cost c is below $I(\tilde{\theta})$, i.e.,

$$G(I(\tilde{\theta})) = G([F_u(\tilde{\theta}) - F_q(\tilde{\theta})]\omega). \tag{14}$$

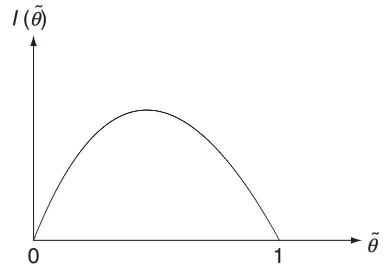

Figure 1 Incentives to invest in skills as a function of the cutoff $\widetilde{\theta}$.

An *equilibrium* of the game is a pair $(\widetilde{\theta}_j^*, \pi_j^*), j \in \{B, W\}$ such that for each j,

$$\widetilde{\theta}_j^* = \widetilde{\theta}(\pi_j^*) \tag{15}$$

$$\pi_j^* = G(I(\widetilde{\theta}_j^*)), \tag{16}$$

where $\widetilde{\theta}(\cdot)$ and $G(I(\cdot))$ are defined by (8) and (14) respectively. Equivalently, we could define the equilibrium of the model as $\pi_j^*, j \in \{B, W\}$, which satisfies:

$$\pi_j^* = G(I(\widetilde{\theta}(\pi_j^*))). \tag{17}$$

From the definition of equilibrium, we see that the only way to rationalize discriminatory outcome for the blacks and whites is when the above equation has multiple solutions.

Existence of multiple equilibria is not always guaranteed and depends on the shape of I and G. This possibility can be proven by construction by fixing all parameters of f_q, f_u, and technology parameters x_q, x_u, ω, and finding an appropriate cost distribution G such that the system (15)–(16) has multiple solutions. Note that since G is a CDF, it is an increasing function of its argument. Therefore, the right-hand side of (16) is a monotone transformation of (13). This means that function (16) must be initially increasing, at least in some range of θ near 0, and subsequently decreasing, at least in some range of θ near 1.

We can find a multitude of functions G that ensure multiple equilibria. For example, assume that all workers have a cost of investment zero or positive, so that $G(0) = 0$. In this case there is always a trivial equilibrium with $\pi = 0, \widetilde{\theta} = 1$. To ensure existence of at least one interior equilibrium, pick $\theta' \in (0, 1)$, and compute π' by inverting (15). Next, compute $I(\theta')$ from (13). If there are a fraction π' of workers with cost less than or equal to $I(\theta')$, then π' is an equilibrium, and there is an infinite number of distributions G that satisfy this condition. Using the same logic, one can construct G functions that are consistent with more than one interior equilibria. This is illustrated in Figure 2, which we drew assuming that there exists some $\widetilde{\theta}$ at which the curve $G(I(\cdot))$ is higher than the inverse of $\widetilde{\theta}(\cdot)$.

When groups select different solutions to Equation (17), they will display different equilibrium human capital investment, employment, and average wages despite having identical fundamentals regarding investment cost and information technology. Thus, Coate and Loury demonstrate that statistical discrimination is a logically consistent

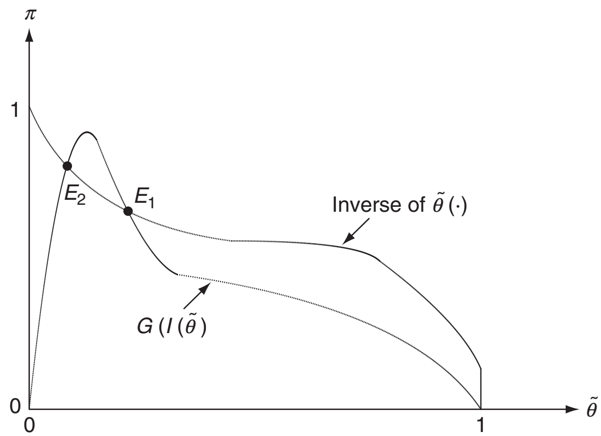

Figure 2 Multiple equilibria in Coate and Loury (1993a).

notion in their model. Discrimination in this model can be viewed as a coordination failure. Equilibria in this model are also Pareto-ranked, as it can be shown that both the workers and the firms would strictly prefer to be in the equilibrium where a higher fraction of workers invests in skills. Group inequality would be eliminated if somehow the blacks and the firms could coordinate on the good equilibrium. Importantly, there is no conflict of interests between whites and blacks concerning the equilibrium selection: if blacks were to coordinate on the better equilibrium, whites would not at all be affected. However, efficiency considerations are somewhat incomplete in this model because wages are set exogenously. We will describe efficiency in equilibrium models of statistical discrimination in more detail in Section 7.

4. DISCRIMINATORY OUTCOMES DUE TO INTER-GROUP INTERACTIONS

In Coate and Loury (1993a), discriminatory outcomes arise in a model where groups could live in separate islands. The privileged group will have no objection whatsoever if the disadvantaged group is able to coordinate themselves into the Pareto dominant equilibrium. In many real-world scenarios, however, we observe conflicts of interest between groups. Models that introduce inter-group interactions in the labor market yield some important insights regarding the potential sources of discrimination. In this section, we describe this literature.

4.1 Discrimination as group specialization

4.1.1 A model with production complementarities and competitive wages

Moro and Norman (2004) relaxed the crucial assumptions guaranteeing group separation in Coate and Loury's model: the linearity of the production technology and the exogeneity of wages. They extended Coate and Loury's framework by assuming a more general technology. In their model output is given by $y(C, S)$, where S is the quantity of workers employed in the simple task, and C is the quantity of *qualified* workers assigned to the complex task; y is strictly quasi-concave, exhibits constant returns to scale and satisfies Inada conditions so that both factors are essential. We use the notation introduced in Section 3.2, and write $x_q(C, S)$ and $x_u(C, S)$ as the marginal products of a *qualified* worker in the complex task, and of any worker employed in the simple task, which now depend on aggregate inputs.

We now characterize the equilibrium in this model. A *Bayesian Nash equilibrium* of the game is a list including the workers' skill investment decision for each cost c, firms task assignment rules, and wage schedules such that every player optimizes against other players' strategy profiles. It can be shown that the optimal task assignment is a threshold rule almost everywhere, where only workers above the threshold $\widetilde{\theta}_j, j = B, W$, are employed in the complex task. Recall that group shares are denoted with λ_j, $j = B, W$. Factor inputs can be computed as follows:

$$S = \sum_{j \in \{B,W\}} \lambda_j \left[\pi_j F_q(\widetilde{\theta}_j) + (1 - \pi_j) F_u(\widetilde{\theta}_j) \right]$$
$$C = \sum_{j \in \{B,W\}} \lambda_j \pi_j \left(1 - F_q(\widetilde{\theta}_j) \right).$$

The thresholds have to be jointly determined for the two groups, because the values of x_q and x_u depend on both groups' assignment rules, given both groups' aggregate investment π_j. The first order conditions are derived from $\max_{\{\widetilde{\theta}_B, \widetilde{\theta}_W\}} y(C, S)$, which are given by:

$$\begin{aligned} &\left[\pi_j f_q(\widetilde{\theta}_j) + (1 - \pi_j) f_u(\widetilde{\theta}_j) \right] x_u(C, S) = \pi_j f_q(\widetilde{\theta}_j) x_q(C, S) \\ &\Rightarrow \frac{\pi_j f_q(\widetilde{\theta}_j)}{\pi_j f_q(\widetilde{\theta}_j) + (1 - \pi_j) f_u(\widetilde{\theta}_j)} = \frac{x_u(C, S)}{x_q(C, S)}, j = B, W \end{aligned} \tag{18}$$

It shows that the input factor ratio C/S is monotonically increasing with the fraction of investors of any group. To see this, note that, if it decreased when π_j increased, then the right-hand side of (18) would decrease. But then the only way to satisfy the first order condition is to decrease $\widetilde{\theta}_j$, because the left-hand side is decreasing in $\widetilde{\theta}_j$ due to the monotone likelihood ratio property assumed for f_q and f_u. However, if both $\widetilde{\theta}_j$ decrease and π_j increase then the factor ratio increases a contradiction.

To understand how this implication affects group incentives to invest in human capital, note that the incentive to invest in Coate and Loury $[F_u(\widetilde{\theta}) - F_q(\widetilde{\theta})]\omega$ may increase or decrease depending only on the value of $\widetilde{\theta}$, because wages are set exogenously. Moro and Norman instead derive wages in equilibrium as the outcome of firms

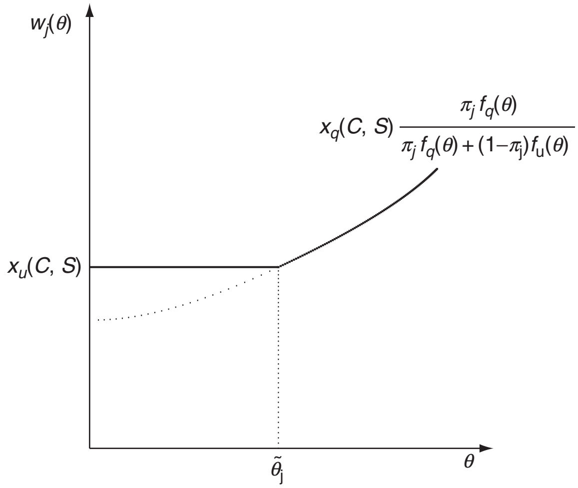

Figure 3 Wage as a function of the signal for group *j*.

competing for workers. It is possible to show that the solution corresponds to wages equal to the expected marginal productivity for almost all $\theta \in [0, 1]$, that is:

$$w_j(\theta) = \begin{cases} x_u(C,S) & \theta < \widetilde{\theta}_j \\ x_q(C,S)\dfrac{\pi_j f_q(\theta)}{\pi_j f_q(\theta) + (1-\pi_j) f_u(\theta)} & \theta \geq \widetilde{\theta}_j \end{cases}. \tag{19}$$

Figure 3 depicts $w_j(\theta)$. Note that the signal value $\widetilde{\theta}_j$ is the one that equates the marginal products in the two tasks, because the term multiplied by $x_q(C, S)$ is the probability that a worker with signal θ is qualified (see equation (4)).

4.1.2 Cross-group effects

We can now compute incentives to invest and indicate them as a function of the vector of investment of the two groups $\boldsymbol{\pi} \equiv (\pi_B, \pi_W)$:

$$I(\boldsymbol{\pi}) = \int_\theta w_j(\theta) f_q(\theta) d\theta - \int_\theta w_j(\theta) f_u(\theta) d\theta.$$

To understand how groups interact, consider the effect on group-B incentives from an increase in π_W. As π_W increases, as noted above, the factor ratio C/S increases. The effect on the marginal product is to increase x_u and decrease x_q. The threshold $\widetilde{\theta}_B$ increases (at the margin, it becomes relatively more convenient to use W workers for the complex task because their likelihood to be qualified increases). This implies that it is more likely for a B worker to be assigned to the simple task (where wages are independent on the signal). Fewer B workers are assigned to the complex task and their wage is a flatter function of the signal than before. Taken together, these observations imply that incentives to invest in human capital decrease when the investment of members of the other group increase.[10]

This result is crucial because it generates incentives for groups to specialize in employment in different jobs. This creates the possibility for asymmetric equilibria to exist even when there is a unique symmetric equilibrium (symmetric equilibria where groups invest in the same proportion are always a possibility).

One asymmetric equilibrium can be constructed by assuming a distribution of investment cost with $G(0) > 0$, that is, by assuming that a fraction $G(0)$ of workers always invest.[11] Assume $\pi_B^* = G(0)$, and that the employers assign all B workers to the simple task. This is optimal if the marginal product of the group-B worker with signal $\theta = 1$ in the complex task is smaller than her marginal product in the simple task:

$$\frac{\pi_B f_q(1)}{\pi_B f_q(1) + (1 - \pi_B) f_u(1)} x_q(C, S) < x_u(C, S) \tag{20}$$

This inequality holds when $G(0) = \pi_B$ is small enough so that the left hand side is small. Note that this is true for any value of input factors C and S, which are not affected by the value of π_B when this inequality holds, because all B workers are in the simple task. To complete the characterization one has to find the equilibrium investment for group W, π_W. However, once group-B workers' behavior is set, the equilibrium level of π_W^* is just the solution of a fixed-point equation in π_W, which by continuity always exists. The equilibrium level of π_W^* must be interior because both factors are essential. The essentiality of both tasks implies that in equilibrium some group-W workers must be employed in the complex task, which implies that incentives to invest are positive for them, and therefore $\pi_W^* > \pi_B^* = G(0)$.

While other equilibria with both groups at an interior solution are possible, it is important to note that such equilibria cannot be interpreted as group-B's failure to coordinate on a better outcome. It is not possible for group-B workers to re-coordinate and invest as white workers do, because when workers of both groups invest in proportion π_W^*, the optimal factor ratio changes and marginal products are no longer consistent with equilibrium.

4.1.3 The effect of group size

Constant returns to scale imply that only *relative* group size matters. In general, analyzing group size effects would mean comparing different sets of equilibria. Not only the analysis becomes more complicated, but also as one parameter such as relative group size changes,

[10] The effect on incentives of group W of an increase in the same group's investment π_W is instead indeterminate, because we also have to take into account the informational externality that acts within groups. When investment increases in one group, the probability of being qualified of all workers from that group increases. This has a beneficial effect on the slope of the increasing portion of the wage function, which may overcome the negative "price" effect on the marginal products of labor we mentioned when we describe the cross-group effects.

[11] With some additional assumptions, it is possible to ensure that the model displays a unique symmetric equilibrium. See Moro and Norman (1996).

some equilibria may disappear and new ones may appear. Therefore, results depend on the details of the equilibrium selection. Intuitively, as the relative size of one group increases and approaches 1, equilibrium investment for this group will approach the values corresponding to the symmetric equilibria of the model (which are equivalent to the equilibrium of a model with only one group). As for the smaller group, depending on the parameterization either lower or higher investment could be consistent with equilibrium.

Nevertheless, we can rely on the simple corner solution constructed in example at the end of the previous section to understand the importance of group size. Because both factors are essential, as discriminated group becomes larger, it becomes more difficult to sustain the extreme type of task segregation implied by the discriminatory equilibrium constructed in the previous section. To see this, note that as the discriminated group becomes larger, the mass of workers employed in the simple task gets larger, and therefore the ratio or marginal products x_u/x_q gets smaller; eventually, the inequality (20) cannot be satisfied and some group-B workers have to be employed in the complex task. Then the incentives to invest in human capital for B workers become strictly positive.

Hence, in a sense, sustaining extreme segregation in equilibrium against large groups may be difficult, rationalizing the existence of institutionalized segregation, such as apartheid in South Africa, where the larger group was segregated into lower paying tasks before the collapse of apartheid. It can also be shown that the incentives for the small group workers to keep the larger group into the segregation-type of equilibrium gets larger the bigger the large group is. The reason is that the larger the mass of workers employed in the simple task is, the higher is the marginal product in the complex job. This increases the incentives to invest for the small group and their benefits from investment.

4.2 Discrimination from search frictions

All theories of statistical discrimination we have described so far are based on information friction in the labor market: race-dependent hiring policies are followed because race is used as a proxy for information about the workers' skills. However, all workers are paid their marginal product and, given skills, color does not play any additional role in explaining racial wage differences once we control for racial differences in their skill investment decisions. That is, there is no "economic discrimination" in the sense of Cain (1986).

Mailath, Samuelson, and Shaked (2000) proposed a model of an integrated labor market and focused on search frictions instead of information friction.[12] As in Moro and Norman (2004), they can derive discriminatory equilibria from a model that displays a unique symmetric equilibrium, but the distinguishing feature of search frictions is that discrimination arises even when employers have perfect information about workers' productivity.

[12] Early examples of statistical discrimination based on a search framework can be found in Verma (1995) and Rosén (1997). Eeckhout (2006) provides a different rationale for inequality arising in a search-matching environment. See Section (5) for more details.

Consider a continuum of firms and workers. All firms are identical, but each worker belongs to either group B or W. Group identity does not directly affect payoffs. For simplicity, suppose that the fraction of group W workers in the population, λ, is equal to 1/2.

All workers are born unskilled, and they make skill investment decisions before entering the labor market. If one acquires skills, he can enter the skilled labor market; otherwise, he enters the unskilled labor market. The crucial difference from the models we have seen so far is that there is no informational friction, that is, *workers' skill investment decisions are observed to the firms*. An individual's skill investment cost $c \geq 0$ is independently drawn from the distribution $G(\cdot)$. Finally, firms and workers die with Poisson rate δ and new firms and workers replace them so that the total populations of both firms and workers are constant. Time is continuous with interest rate r.

Each firm can hire at most one worker. If a firm employs a skilled worker, regardless of his color, a flow surplus of $x > 0$ is generated; the flow surplus from hiring an unskilled labor is 0.

Search frictions and wage determination. Vacant firms, meaning firms without an employee, and unemployed workers match through searches. Searches are assumed costless for both the firms and the workers. Given the assumption that the surplus for a firm from hiring an unskilled worker is 0, firms will only search for skilled workers. Firms make a key decision of whether to search either groups, or only one group. Suppose that a firm searches for workers of both groups, and suppose that the proportion of the skilled workers in the population is H_I and the unemployment rate of skilled workers is ρ_I, then the process describing meetings between unemployed skilled workers and the searching firm follows a Poisson process with meeting rate $\gamma_F \rho_I H_I$ where the parameter γ_F captures the intensity of firm search. If instead, the firm searches only white workers with intensity γ_F, then the meeting rate between the firm and the white skilled workers is given by $2\gamma_F \rho_I H_I$. Unemployed skilled workers simultaneously search for vacant firms with intensity γ_I and the meetings generated by workers search follow a Poisson process with rate $\gamma_I \rho_F$ where ρ_F is the vacancy rate of the firms. When an unemployed worker and a vacant firm match, they bargain over the wage with one of them randomly drawn to propose a take-it-or-leave-it offer.

Symmetric Steady State Equilibrium. We first characterize the symmetric steady state equilibrium in which firms do not pay any attention to the workers' color so we can treat the workers as a single population. We use subscript I to denote worker related variables in this section. Let V_I denote the value of skills to an individual in equilibrium. Since an individual will invest in skills only if his skill investment cost c is less than V_I, the fraction of skilled workers in the population will be $G(V_I)$. Let H_I be the proportion of skilled workers in the population in the steady state. We must have:

$$H_I = G(V_I) \tag{21}$$

in the steady state. The steady state condition for vacancies ρ_F is given by:

$$2\delta(1-\rho_F) = \rho_F \rho_I H_I(\gamma_I + \gamma_F). \tag{22}$$

In (22), the left hand side represents the rate of vacancy creation because $1 - \rho_F$ is the fraction of firms which are currently occupied, and at the rate 2δ either a worker dies, creating a vacancy at a previously occupied firm, or an occupied firm dies, and is replaced by a new vacant firm. The right hand side is the rate of vacancy destruction because of matches formed due to worker or firm searches. Similarly, the steady state condition for unemployment rate of the skilled worker ρ_I is given by:

$$2\delta(1-\rho_I) = \rho_F \rho_I(\gamma_I + \gamma_F). \tag{23}$$

Finally, we need to derive V_I. Let ω be the expected flow payoff of an employed worker and Z_I be the steady-state value of an employed skilled worker. First, familiar results from dynamic programming give us:

$$(r+2\delta)\ Z_I = \omega + \delta V_I,$$

where the left hand side $(r + 2\delta)\ Z_I$ can be interpreted as the properly normalized flow payoff of an employed worker, which is exactly equal to the wage ω plus, with probability δ, the worker obtains the expected present value of being returned to the unemployment pool by surviving a firm death, V_I. Similarly, when a skilled worker is unemployed, his value V_I is related to Z_I as follows:

$$[\rho_F(\gamma_F+\gamma_I) + r + \delta]\ V_I = \rho_F(\gamma_F + \gamma_I)\ Z_I.$$

On the firm side, let ϕ be the expected flow payoff to an occupied firm, V_F be the steady state value of a vacant firm, and Z_F be the steady state value of a firm who is currently employing a skilled worker. Since $\omega + \phi = x$, the total flow surplus, we know that the total surplus when a vacant firm and an unemployed worker match, denoted by S, must satisfy:

$$(r+2\delta)\ S = x + \delta(V_I + V_F).$$

Since the firm and the worker divide the surplus from the relationship relative to the status quo, given by $S - V_F - V_I$, via Nash bargaining, we have:

$$Z_I = V_I + \frac{1}{2}(S - V_F - V_I),$$

$$Z_F = V_F + \frac{1}{2}(S - V_F - V_I).$$

Thus, we can obtain:

$$V_I = \frac{\rho_F(\gamma_F+\gamma_I)x}{(r+\delta)[(\rho_F+\rho_I H_I)(\gamma_F+\gamma_I) + 2(r+2\delta)]}, \tag{24}$$

$$V_F = \frac{\rho_I H_I(\gamma_F + \gamma_I)x}{(r+\delta)[(\rho_F + \rho_I H_I)(\gamma_F + \gamma_I) + 2(r + 2\delta)]} \tag{25}$$

A symmetric steady state is a list $(H_I, \rho_I, \rho_F, V_I, V_F)$ satisfying the steady state conditions (21)–(25). A symmetric steady state is a symmetric equilibrium if the postulated search behavior of the firms, i.e., each firm searches both colors of workers, is optimal. Obviously, since the two groups of workers are behaving identically, any symmetric steady state will indeed be a symmetric equilibrium. With some algebra, Mailath, Shaked, and Samuelson showed that a symmetric equilibrium exists and is unique.

Asymmetric Equilibrium. Now consider the asymmetric equilibrium in which firms search *only white workers*. Under the postulated search behavior of the firms, skilled black workers can be matched to firms only through the worker searches, but the skilled white workers can be matched to firms both through the searches initiated by the workers and the firms. Now first consider the steady state conditions for the postulated asymmetric equilibrium. In this section, we use subscript W and B to denote group-W and group-B related variables respectively.

Let H_W and H_B denote the fraction of skilled workers among white and black population respectively, and let V_W and V_B denote the value of skill for white and black workers respectively. As in the symmetric equilibrium case, the skilled worker steady state conditions are:

$$H_W = \frac{G(V_W)}{2}, \tag{26}$$

$$H_B = \frac{G(V_B)}{2}. \tag{27}$$

Likewise, the vacancies steady state condition will now read:

$$2\delta(1-\rho_F) = 2\rho_F\gamma_F H_W \rho_W + (\rho_W H_W + \rho_B H_B)\gamma_I \rho_F. \tag{28}$$

The white and black unemployment rate steady state conditions are:

$$2\delta(1-\rho_W) = \rho_W \rho_F(\gamma_I + 2\gamma_F). \tag{29}$$

$$2\delta(1-\rho_B) = \rho_B \rho_F \gamma_I. \tag{30}$$

Now we characterize the relevant value functions in an asymmetric steady state. Let ω_j, $j \in \{B, W\}$, be the expected wage of a skilled worker with race j, Z_j be the present value of a race-j employed skilled worker, V_F be the present value of a vacant firm, and $Z_{F,j}$ be the present value of a firm matched with a race-j skilled worker. We have the following relationships:

$$
\begin{aligned}
(r+2\delta)Z_j &= \omega_j + \delta V_j, j \in \{B, W\}, \\
(r+2\delta)Z_{F,j} &= \phi_j + \delta V_F, j \in \{B, W\}, \\
(\rho_F\gamma_I + r + \delta)V_B &= \rho_F\gamma_I Z_B, \\
V_B &= \rho_F(\gamma_I + 2\gamma_F)Z_W,
\end{aligned}
$$

Derivations similar to those for the symmetric steady state yield the following value functions in a white asymmetric steady state:

$$
V_F = \frac{x}{(r+\delta)\Delta}\left\{ \begin{array}{c} (2\gamma_F + \gamma_I)\rho_F\gamma_I(\rho_W H_W + \rho_B H_B) \\ +2(r+2\delta)[(2\gamma_F+\gamma_I)\rho_W H_W + \gamma_I\rho_B H_B] \end{array} \right\}, \tag{31}
$$

$$
V_B = \frac{\rho_F\gamma_I[2(r+2\delta) + (2\gamma_F+\gamma_I)\rho_F]x}{(r+\delta)\Delta}, \tag{32}
$$

$$
V_W = \frac{\rho_F(2\gamma_F+\gamma_I)[2(r+2\delta) + \rho_F\gamma_I]x}{(r+\delta)\Delta}, \tag{33}
$$

where $x = \omega_j + \phi_j$, $j \in \{B, W\}$, is the total surplus, and:

$$
\begin{aligned}
\Delta \equiv\ & 2(r+2\delta)[(2\gamma_F+\gamma_I)(\rho_F + \rho_W H_W) + \gamma_I(\rho_F + \rho_B H_B) + 2(r+2\delta)] \\
& + \rho_F\gamma_I(2\gamma_F+\gamma_I)(\rho_F + \rho_W H_W + \rho_B H_B).
\end{aligned}
$$

A *white asymmetric steady state* is a list (H_W, ρ_W, V_W, H_B, ρ_B, V_B, ρ_F, V_F) such that the balance equations (26)–(30) and the value functions (31)–(33) hold. It can be verified that in a white asymmetric steady state, black workers face a less attractive value of entering the skilled labor market than do white workers ($V_B < V_W$), and thus fewer black workers than white workers acquire skills ($H_B < H_W$). Black workers thus are at a disadvantage when bargaining with firms and, as a result, firms obtain a larger surplus from black workers ($\omega_B < \omega_W$ and $\phi_B > \phi_W$). Given this pattern of surplus sharing, a vacant firm would prefer to hire a black skilled worker than a white skilled worker ($Z_{F,B} > Z_{F,W}$). Moreover, since it is postulated that firms are only searching for white skilled workers, it must be the case that unemployment rate is higher among blacks than among whites ($\rho_B > \rho_W$).

However, in order for the postulated white asymmetric steady state to be consistent with equilibrium, the firms must find it optimal to only search the white workers. Let $V_F\ (B|W)$ ($V_F\ (BW|W)$, respectively) be the value of a firm searching only black workers (searching both black and white workers, respectively) if the other firms are all searching only the white workers. It can be shown that they are respectively given by:

$$
V_F(B|W) = \frac{\gamma_I\rho_W H_W Z_{F,W} + (2\gamma_F+\gamma_I)\rho_B H_B Z_{F,B}}{\gamma_I\rho_W H_W + (2\gamma_F+\gamma_I)\rho_B H_B + r + \delta}, \tag{34}
$$

$$V_F(BW|W) = \frac{(\gamma_I + \gamma_F)(\rho_W H_W Z_{F,W} + \rho_B H_B Z_{F,B})}{(\gamma_I + \gamma_F)(\rho_W H_W + \rho_B H_B) + r + \delta}. \tag{35}$$

The condition for white asymmetric steady state equilibrium is:

$$V_F \geq \max\{V_F(B|W), V_F(BW|W)\}. \tag{36}$$

Examining the expressions for V_F, $V_F(B|W)$ and $V_F(BW|W)$ as given by (31), (34) and (35), we can see that (36) can be true only if $\rho_B H_B < \rho_W H_W$ in a white asymmetric equilibrium. Since we already know that $\rho_B > \rho_W$ in the asymmetric steady state, it thus must be the case that $H_B < H_W$. That is, to be optimal for the firms to only search for white workers in the white asymmetric equilibrium, there must be a sufficiently low fraction of skilled black workers. That is, the postulated discriminatory search behavior of the firms in favor of whites must generate a sufficiently strong *supply side response* on the part of workers in their skill investment decisions in order for the firms' search behavior to be optimal. The intuition is quite simple: In order for the firms not to search for blacks, and knowing that in equilibrium the wages for black skilled workers are lower, it must be the case that there are a lot fewer black skilled workers in order for the trade-off between a larger surplus from each hired black worker and a smaller probability of finding such worker to be in favor of not searching blacks.

Mailath, Samuelson, and Shaked (2000) show that a sufficient condition for a white asymmetric equilibrium is that when firms' search intensity γ_F is sufficiently large relative to that of the search intensity of the workers γ_I. The intuition for this result is as follows: when firms' searches are responsible for a sufficiently large fraction of the contacts between firms and workers, a decision by the firms not to search the black workers will almost ensure that skilled black workers would not find employment; thus, depressing their incentives to acquire skills, which in turn justifying the firms' decision not to search the black workers. Therefore, this paper shows how search friction might generate group inequality even when employers have perfect information about their workers and would strictly prefer to hire workers from the discriminated group. Holden and Rosén (2009) show in a similar framework that the existence of prejudiced employers may also make it more profitable for nondiscriminatory employers to discriminate.

4.3 Endogenous group formation

The models presented so far assume that individuals' group identities are exogenous. In some situations, group identity is not as immutable as one's skin color or gender, but is defined by characteristics that are more amenable to change, albeit at costs. Fang (2001) presents a model of discrimination with endogenous group formation, where he showed that endogenous group formation and discrimination can in fact coexist, and the resulting market segmentation in the discriminatory equilibrium may lead to welfare improvement. Relative to Coate and Loury (1993a), Fang's model keeps their linear production technology, but endogenizes group identity choices; in addition, wages are set endogenously *à la* Moro and Norman (2003a) (see Section 4.1).

Benchmark Model with No Group Choice. The benchmark is a model without endogenous group choices. There are two (or more) firms, indexed by $i = 1, 2$. They both have a traditional (old) and a new technology at their disposal. Every worker can produce 1 unit of output with the traditional technology. Workers with some requisite skills can produce $x_q > 1$ units of outputs with the new technology, but those without the skills will produce 0. We assume that the firms are risk neutral and maximize expected profits.

There is a continuum of workers of unit mass in the economy. Workers are heterogeneous in their costs of acquiring the requisite skills for the new technology. Suppose for simplicity that a worker is either a *low cost type* whose skill acquisition cost is c_l or a *high cost type* with cost c_h where $0 < c_l < c_h$. The fractions of low cost and high cost workers are λ_l and λ_h respectively with $\lambda_l + \lambda_h = 1$. A worker's cost type is her private information. It is assumed that the workers are risk neutral and that they do not directly care about the technology to which they are assigned.

To dramatize the market failure caused by informational free riding, suppose that it is socially optimal for every worker to invest in skills and use the new technology, i.e., $x_q - c_h > 1$.

The timing of the game and the strategies of the players are as follows. First, workers, observing their cost realization $c \in \{c_l, c_h\}$, decide whether to invest in skills, e: $\{c_h, c_l\} \to \{e_q, e_u\}$. Firms do not perfectly observe a worker's investment decision, instead they observe in the second stage a signal $\theta \in [0, 1]$ about each worker. The signal θ is distributed according to probability density function f_q for qualified workers and f_u for unqualified ones. We assume that $f_q(\cdot)/f_u(\cdot)$ is strictly increasing in θ. In the third stage, the firms compete in the labor market for workers by simultaneously announcing wage schedules as functions of the test signal θ. A pure action of firm i at this stage is a mapping $w_i : [0, 1] \to \mathfrak{R}_+$. Workers then decide for which firm to work after observe wage schedules w_1 and w_2. Finally, each firm allocates its available workers between the old and new technologies using an assignment rule which is a mapping $t_i : [0, 1] \to \{0, 1\}$, where $t_i(\theta) = 1$ (respectively, 0) means that firm i assigns all workers with signal θ to the new (respectively, old) technology.

A *Bayesian Nash equilibrium* of the game is a list including the workers' skill investment decision profile e and offer acceptance rules, and the firms' wage schedules and technology assignment rules $\{w_i(\cdot), t_i(\cdot)\}$ such that every player optimizes against other players' strategy profiles. Wages in equilibrium must be equal to workers' expected marginal product for almost all $\theta \in [0, 1]$, as in equation (19):

$$w_1(\theta) = w_2(\theta) = w(\pi, \theta) \equiv \max\left\{1, \frac{\pi f_q(\theta)}{\pi f_q(\theta) + (1-\pi) f_u(\theta)} x_q\right\}; \qquad (37)$$

and the firms' equilibrium assignment rule must be $t_1(\theta) = t_2(\theta) \equiv t(\theta)$, where $t(\theta) = 1$ if for almost all $\theta \in [0, 1]$:

$$\frac{\pi f_q(\theta)}{\pi f_q(\theta) + (1-\pi) f_u(\theta)} x_q \geq 1.$$

To analyze workers' skill investment decisions in Stage 1, note that the *private benefit* of skill investment when a fraction π of the population is skilled is:

$$I(\pi) = \int_0^1 w(\pi, \theta)[f_q(\theta) - f_u(\theta)] d\theta.$$

Because the private benefit is a function of π, there is *informational free riding*. In fact, the informational free riding problem may lead to $\pi = 0$ being the unique equilibrium outcome. Specifically, define Π_l and Π_h to be the sets of values of π that will respectively induce low and high cost type workers to invest in the skills, that is, $\Pi_l \equiv \{\pi \in [0, 1]: I(\pi) \geq c_l\}$; $\Pi_h \equiv \{\pi \in [0, 1]: I(\pi) \geq c_h\}$. Then it can be shown that any economy where $\Pi_l \neq \emptyset$ and $\min \Pi_l > \lambda_l$; but $\Pi_h = \emptyset$ will have a unique equilibrium with $\pi = 0$. The intuition is analogous to a domino effect: $\Pi_h = \emptyset$ implies that type-c_h workers will never invest in skills, but the presence of the high cost types dilutes the benefit of skill investment for type-c_l types.

Endogenous Group Choices and Discriminatory Equilibrium. Now suppose there is an activity A that workers can undertake. Let $V \in \mathfrak{R}$ be a worker's utility (or disutility if negative) in monetary terms from activity A. Therefore, each worker now has two private characteristics (c, V). Let $H(V|c)$ denote the cumulative distribution of V conditional on the skill acquisition cost c. Importantly, assume that whether a worker undertakes activity A is *observable* to firms. The defining characteristic of a cultural activity is that it is *a priori* completely irrelevant to other economic fundamentals, which is captured by two assumptions: (1). $H(V|c_l) = H(V|c_h) \equiv H(V)$, where H is continuous and strictly increasing in V with support $[\underline{V}, \bar{V}] \subset \mathfrak{R}$; (2) A worker's test signal, and her qualification for the new technology are not affected by whether she undertakes activity A. The game is expanded to include one additional stage where a worker of type (c, V) chooses $j \in \{A, B\}$, where $j = A$ means that she undertakes activity A and $j = B$ that she does not. She derives from activity A (dis)utility V if she chooses $j = A$, and zero utility otherwise. Write the activity choice profile as $g: \{c_l, c_h\} \times [\underline{V}, \bar{V}] \to \{A, B\}$. Workers who choose A will be called A-*workers*, and those who choose B, B-*workers*.

Because of the *a priori* irrelevance of activity A we can suitably augment the equilibrium decision rules of the basic model, and obtain an equilibrium of the augmented model in which activity A plays no role in the firms' wage offer schedules and technology assignments. The activity and skill acquisition choices in this type of equilibrium,

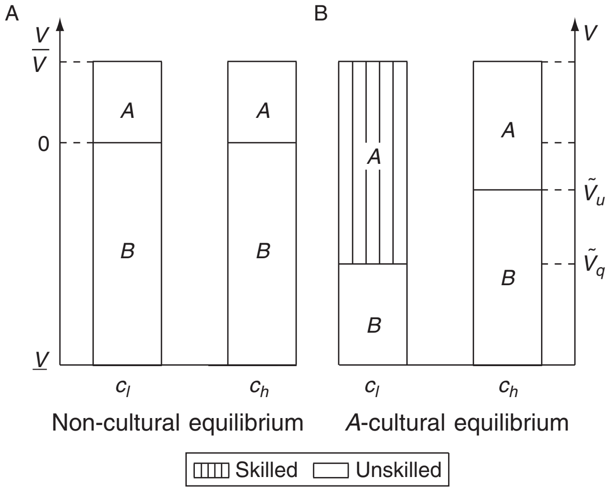

Figure 4 Activity and Skill Acquisition Choices: Fang (2001).

dubbed "non-cultural equilibrium," are pictured in Figure (4a). It is obvious that in the non-cultural equilibrium, no workers are skilled; hence, the new technology is not adopted.

The introduction of the observable activity A allows the firms potential to offer wage schedules and technology assignment rules contingent on whether activity A is undertaken. If firms do use this type of contingent wage schedules then workers may undertake activity A for instrumental reasons. If A-workers are preferentially treated (in a manner to be made precise below), then some workers who intrinsically dislike activity A may choose A to get the preferential treatment. Of course, in equilibrium it must be rational for firms to give preferential treatment to A-workers.

An *A-cultural equilibrium* is defined to be a Bayesian Nash equilibrium of the augmented model in which a positive mass of A-workers are assigned to the new technology, while all B-workers are assigned to the old technology. Now consider an A-cultural equilibrium. Since B-workers are never assigned to the new technology, in this equilibrium the fraction of the skilled among B-workers, denoted by π_B, must be zero. Furthermore, in order for some positive fraction of A-workers to be assigned to the new technology, the proportion of the skilled among A-workers, denoted by π_A, must belong to the set Π_l. An A-cultural equilibrium exists if for some value $\pi_A \in \pi_l$, the population will self-select the activity choices such that the fraction of c_l types among A-workers is exactly π_A.

As before, workers will still be paid their expected productivity. Therefore firm i's sequentially rational wage offer schedule to B-workers, w_i^B, is:

$$w_1^B(\theta) = w_2^B(\theta) = w(0,\theta) = 1 \text{ for all } \theta \in [0,1].$$

Suppose that the proportion of the skilled among A-workers is π_A. Then firm i's equilibrium wage schedule to A-workers, w_i^A, is:

$$w_1^A(\theta) = w_2^A(\theta) = w(\pi_A, \theta).$$

For every π_A, the expected wage of a skilled A-worker is $W_q^A(\pi_A) = \int_0^1 w(\pi_A, \theta) f_q(\theta)\, d\theta$, and that of an unskilled A-worker is $W_u^A(\pi_A) = \int_0^1 w(\pi_A, \theta) f_u(\theta)\, d\theta$. We can prove, by simple revealed preference arguments that the activity and skill-acquisition choice profiles under an A-cultural equilibrium, where the proportion of the skilled among A-workers is π_A, must be:

$$e(c, V) = \begin{cases} e_q & \text{if } c = c_l,\ V \geq 1 + c_l - W_q^A(\pi_A) \\ e_u & \text{otherwise} \end{cases}$$

$$g(c, V) = \begin{cases} A & \text{if } c = c_l,\ V \geq 1 + c_l - W_q^A(\pi_A) \\ A & \text{if } c = c_h,\ V \geq 1 - W_u^A(\pi_A) \\ B & \text{otherwise.} \end{cases}$$

Figure (4b) depicts the activity and skill acquisition choices in an A-cultural equilibrium where we have defined $\widetilde{V}_q(\pi_A) = 1 + c_l - W_q^A(\pi_A)$ and $\widetilde{V}_u(\pi_A) = 1 - W_u^A(\pi_A)$ as the threshold disutility values that respectively a skilled and an unskilled worker are willing to incur to be a member of the elites. Note that $W_q^A(\pi_A) - W_u^A(\pi_A) \geq c_l$ because $\pi_A \in \Pi_l$. Since $W_u^A(\pi_A) \gg 1$ whenever there is a positive mass of A-workers assigned to the new technology, we have:

$$\widetilde{V}_q(\pi_A) \ll \widetilde{V}_u(\pi_A) \ll 0. \tag{38}$$

Inequality (38) establishes that in a cultural equilibrium, a single-crossing property of the cultural activity is *endogenously* generated. More specifically, let us denote the *net* benefit to undertake activity A for a skilled and an unskilled worker with the same utility type V respectively by $B(e_q, V; \pi_A) = V - \widetilde{V}_q(\pi_A)$ and $B(e_u, V; \pi_A) = V - \widetilde{V}_u(\pi_A)$. Inequality (38) yields that $B\,(e_q, V; \pi_A) > B\,(e_u, V; \pi_A)$ for every type V. In other words, in any A-cultural equilibrium, a skilled worker is more willing than an unskilled one to endure disutility from activity A to be elite, which in turn justifies A-workers as elites. Undertaking activity A becomes a signaling instrument for skilled workers due to the endogenously generated single crossing property.

Fang (2001) provided the necessary and sufficient condition for the existence of A-cultural equilibria. For any $\pi_A \in \Pi_l$, define the proportion of the skilled among A-workers by a mapping $\Psi : [0, 1] \to [0, 1]$ given by:

$$\Psi(\pi_A) = \begin{cases} \dfrac{\lambda_l(1 - H(\widetilde{V}_q(\pi_A)))}{\lambda_l(1 - H(\widetilde{V}_q(\pi_A))) + \lambda_h(1 - H(\widetilde{V}_u(\pi_A)))} & \text{if } \pi_A \in \Pi_l \\ 0 & \text{otherwise} \end{cases}$$

where the numerator of the fraction is the total mass of skilled A-workers (see the shaded area in Figure 4b) and the denominator is the total mass of A-workers (the area marked "A" in Figure 4b). Every fixed-point of the mapping Ψ will correspond to an A-cultural equilibrium. Notice that by segmenting the labor market into A-workers and B-workers (by whether workers undertake the activity A,) it allows A-workers' skill investment choices depend only on the firms' perception of the proportion of the skilled among A-workers, instead of the firm's perception for the whole population as in the benchmark model. Let $\Delta \equiv \max_{\pi_A \in \Pi_1}[\Psi(\pi_A) - \pi_A]$ be the maximal difference between the function Ψ and the identity map. The necessary and sufficient condition for the existence of at least one A-cultural equilibrium is $\Delta \geq 0$.

Welfare. In a cultural equilibrium, the new technology is adopted by a positive mass of workers. In the mean time, some workers are enduring the disutility of activity A in order to be members of the elites. However, B-workers are exactly as well off as they are in the non-cultural equilibrium. By revealed preferences, A-workers are strictly better off than they are in the non-cultural equilibrium. Thus, any cultural equilibrium Pareto dominates the non-cultural equilibrium.

4.4 Group interactions from peer effects

An alternative source of cross-group interactions is studied by Chaudhuri and Sethi (2008), who extended the standard Coate and Loury's framework assuming that the distribution of the cost of investment in human capital, G, is a function of the mean peer group skill level s, computed as follows:

$$s_j = \eta\pi_j + (1-\eta)\bar{\pi}, j = B, W$$

where $\bar{\pi}$ is the fraction of skilled workers in the whole population and $\eta \in [0, 1]$ measures the level of segregation in the society. Positive spillover in human capital across groups is reflected in the assumption that G is increasing in s_j. Although G is the same across groups, the distribution of the cost of acquiring human capital for a given group is endogenous in this model, and may be different across groups if groups experience different levels of peer quality.

This parameterization allows the investigation of the relationship between integration and discrimination. Chaudhuri and Sethi show that integration may make it impossible to sustain negative stereotypes in equilibrium. To understand the intuition behind the main result, assume that when groups are completely segregated they coordinate on different equilibria. As integration increases, the peer group effect increases the cost of investment for the group with high investment and decreases the cost of investment for the other group; hence, the direct effect is to equalize the fractions of people that invest. Inequality may persist in equilibrium, but under some conditions, if integration is strong enough multiplicity of equilibria disappears and groups acquire the same level of human capital.

5. DYNAMIC MODELS OF DISCRIMINATION

The literature on the dynamic evolutions of discrimination is relatively sparse. Antonovics (2006) considers a dynamic model of statistical discrimination that accounts for intergenerational income mobility. She shows that when income is transmitted across generations through parental investments in the human capital of children, statistical discrimination can lead racial groups with low endowments of human capital to become trapped in inferior stationary equilibria. Fryer (2007) considers a dynamic extension of the Coate and Loury model, more specifically the example that Coate and Loury used to illustrate the potential for patronizing equilibrium with affirmative action as described in Section 6.2.2, by introducing an additional promotion stage after workers are hired. He uses the extension to ask how initial adversity in the hiring stage will affect the subsequent promotions for those minorities who are able to be assigned a job in the firm. The intuition he formalizes in the model can be termed as a possibility of "belief flipping." Specifically, suppose that an employer has negative stereotypes about a particular group, say the minorities, and discriminates against them in the initial hiring practice, relative to another group, say the majorities, for whom the employer has more stereotypes that are positive. Then, conditional on being hired, the minority workers within the firm may be more talented than the majority workers because they were held to a more exacting standard in the initial hiring. As a result, minorities who are hired in the firm may be more likely to be promoted. Fryer's (2007) analysis provides a set of sufficient conditions for the "belief flipping" phenomenon to arise.[13]

Blume (2006) presents an interesting dynamic analysis of statistical discrimination using ideas from evolutionary game theory. This paper adds a learning dynamic to a simplified version of Coate and Loury's static equilibrium model of statistical discrimination. The assignment of workers to firms and the opportunity for firms to experiment generate a random data process from which firms learn about the underlying proportions of skilled workers in the population. Under two plausible, but exogenously specified learning dynamics, long-run stable patterns of discrimination that appear in the data process can be characterized and related to the equilibria of the static model. Blume (2006) shows that long-run patterns of discrimination can be identified with particular equilibria. Although different patterns corresponding to different equilibria are possible, generically only one will be salient for any given specification of parameters.

Blume's (2006) dynamic model is cast in a discrete time setting where in each period, a certain measure of new workers are born and they will have to make unobservable skill investment decisions. A drawback of the discrete time setup is that there

[13] The flipping of the effect of race on the initial hiring probability and subsequent promotion probability may be used as a basis to empirically distinguish statistical discrimination from taste-based discrimination. Altonji and Pierret (2001) proposed and implemented a test of statistical discrimination based on the effect of race on worker wages over time with employer learning.

will be potential multiple equilibria in the skill investment decisions within each cohort due to coordination failure. Levin (2009) avoids this complication by considering a continuous time model where in any instant only one new worker arrives with some probability, thus avoiding the issue of equilibrium multiplicity resulting from coordination problems. As a result, the evolution of the fraction of skilled workers in Levin (2009) is consistent with the optimal behavior of the individuals. He showed that statistical discrimination equilibrium can be persistent even if policies are enacted to improve access to resources for the disadvantaged minorities.

Eeckhout (2006) provides an alternative theory of discrimination based on a search and matching model of a marriage market. This paper generates endogenous segregation in a dynamic environment where partners randomly match to play a repeated prisoner's dilemma game.[14] In this setup, the driving force behind inequality is the use of race as a public randomization device. When cooperation is expected from same-match partners, segregation outcomes might Pareto-dominate color-blind outcomes. Due to random matching, mixed matches always occur in equilibrium, and there may be less cooperation in mixed matches than in same-color matches, but mixed matches may be of shorter duration.

6. AFFIRMATIVE ACTION

6.1 A brief historical background

Affirmative action policies were developed during the 1960s and 1970s in two phases that embodied conflicting traditions of government regulations.[15] The first phase, culminating in the Civil Rights Act of 1964 and the Voting Rights Act of 1965, was shaped by the presidency and the Congress and emphasized nondiscrimination under a "race-blind Constitution." The second phase, shaped primarily by federal agencies and courts, witnessed a shift toward minority preferences during the Nixon administration. The development of two new agencies created to enforce the Civil Rights Act, the Equal Employment Opportunity Commission under Title VII and the Office of Federal Contract Compliance under Title VI of the Civil Rights Act, demonstrates the tensions between the two regulatory traditions and the evolution of federal policy from non-discrimination to minority preferences under the rubric of affirmative action. The results have strengthened the economic and political base of the civil rights coalition while weakening its moral claims in public opinion.

The main goals of the Civil Rights Act of 1964 were "the destruction of legal segregation in the South and a sharp acceleration in the drive for equal rights for women". Title VII, known as the Fair Employment Commission Title or FEPC Title, of the

[14] Fang and Loury (2005a, 2005b) explored a theory of dysfunctional collective identity in a repeated risk sharing game.

[15] See Holzer and Neumark (2000) for a more detailed historical and institutional background of affirmative action's in the U.S.

Civil Rights Act would create the Equal Employment Opportunity Commission (EEOC) to police job discrimination in commerce and industry with the intention to destroy the segregated political economy of the South and enforce nondiscrimination throughout the nation. Title VI of the Act, known as the Contract Compliance Title, "prohibits discrimination in programs receiving funds from federal grants, loans or contracts." The authority to cancel the contracts of failed performers and ban the contractors from future contract work backed contract compliance. The specter of bureaucrats telling businesses whom to hire under Title VII was raised during the congressional debates prior to the passage of the Civil Rights Act. The Senate majority leader of the time, Hubert Humphrey, promised to eat his hat if the civil rights bill ever led to racial preferences. President Lyndon Johnson signed the Civil Rights Act of 1964 into law on 2 July.

But Title VI of the Civil Rights Act of 1964 was the sleeper that led to affirmative action policies. In September 1965, President Johnson issued Executive Order 11246. This order intended to create new enforcement agencies to implement Title VI in the Civil Rights Act, and it repeated nondiscrimination. The Office of Contract Compliance (OFCC) was established by the Labor Department to implement Executive Order 11246. It designed a model of contract compliance based on a metropolitan Philadelphia plan, which required that building contractors submit "pre-award" hiring schedules listing the number of minorities to be hired, with the ultimate goal to make the proportion of blacks in each trade equal to their proportion of metropolitan Philadelphia's workforce (30%). This Philadelphia plan was ruled in November 1968 to violate federal contract law. Nevertheless, in 1971 under the Nixon administration, the Supreme Court affirmed that the minority preferences of the Philadelphia did not violate the Civil Rights Act. The EEOC, in charge of the implementation of Title VII, followed a similar strategy, issuing guidelines to employers to use statistical proportionality in employee testing. In 1972, Congress extended the EEOC's jurisdiction to state and local governments and educational institutions (which were exempt in 1964). Affirmative action became the model of federal hiring practices.

The original rationale for affirmative action was to right the historical wrong of institutional racism and stressed its temporary nature. In 1978, in *Regents of the University of California vs. Bakke*, Supreme Court justice Harry Blackmun was apologetic about supporting a government policy of racial exclusion: "I yield to no one in my earnest hope that the time will come when an affirmative action program is unnecessary and is, in truth, only a relic of the past." He expressed the hope that it is a stage of transitional inequality and "within a decade at most, American society must and will reach a stage of maturity where acting along this line is no longer necessary." Twenty-five years later, however, in her opinion on the case *Grutter vs. Bollinger*, justice Sandra Day O'Connor repeated a similar rhetoric: "The Court expects that 25 years from now, the use of racial preferences will no longer be necessary to further the interest approved today."

6.2 Affirmative action in Coate and Loury (1993a)

Coate and Loury (1993a) analyzed how affirmative action in the form of an employment quota may affect the incentives to invest in skills for both groups and the equilibrium of the model. In particular, it highlights a potential perverse effect of affirmative action: in the so-called "patronizing equilibrium," the incentives to invest in skills by the group A workers – the group that the affirmative action policy is supposed to help, may be reduced in the equilibrium with affirmative action relative to that without affirmative action.

6.2.1 Modeling affirmative action

Coate and Loury (1993a) modeled affirmative action as an employment quota. Specifically, the affirmative action policy requires that the proportion of group B workers on the complex task (which pays a higher wage in their model) be equal to the proportion of group B workers in the population. Recall from Section 3.2, the proportion of white workers in the population is $\lambda \in (0, 1)$. For expositional simplicity, we write $\lambda_W = \lambda$ and $\lambda_B = 1 - \lambda$ below.

Suppose that the proportions of skilled workers are respectively π_B and π_W among groups B and W. Let

$$\rho(\widetilde{\theta}, \pi) \equiv \pi[1 - F_q(\widetilde{\theta})] + (1 - \pi)[1 - F_u(\widetilde{\theta})]$$

be the probability that the firms will assign a randomly selected worker from a group where a fraction π invests in skills to the complex task if the firms use $\widetilde{\theta}$ as the assignment threshold. Now we can write firms' task assignment problem under the employment quota as:

$$\max_{\{\widetilde{\theta}_W, \widetilde{\theta}_B\}} \sum \lambda_j \{\pi_j[1 - F_q(\widetilde{\theta}_j)]x_q - (1 - \pi_j)[1 - F_u(\widetilde{\theta}_j)]x_u\} \tag{39}$$

$$\text{s.t. } \rho(\widetilde{\theta}_W, \pi_W) = \rho(\widetilde{\theta}_B, \pi_B) \tag{40}$$

where in the affirmative action employment quota constraint (40), the left and right hand sides are respectively the probabilities that a random White and Black worker will be assigned to the complex task. Note that when these probabilities are equalized, the fraction of blacks assigned to the complex task will indeed exactly match the fraction of blacks in the population, as stipulated by the employment quota.[16]

An *equilibrium under affirmative action* is a pair of beliefs $(\pi^*{}_B, \pi^*{}_W)$ and cutoffs $(\widetilde{\theta}^*_B, \widetilde{\theta}^*_W)$ such that: (1) $(\widetilde{\theta}^*_B, \widetilde{\theta}^*_W)$ solves problem (39) given $(\pi^*{}_B, \pi^*{}_W)$; (2) $\pi^*_j = G(I(\widetilde{\theta}^*_j))$ for $j = B, W$.

The ideal for an affirmative action policy is to ensure that all equilibria under affirmative action entail homogeneous beliefs by the firms about the investment behavior of the workers from the two groups and lead to a result of race-neutral task assignment

[16] Assuming a law of large numbers holds in this setup.

decisions. The negative stereotypes of the firms regarding the discriminated against group will be eliminated by the affirmative action policy if firms hold homogeneous beliefs.

Coate and Loury (1993a) provide a sufficient condition on the primitives, albeit rather difficult to interpret, for the above ideal of affirmative action to be realized. Let:

$$\hat{\rho}(\widetilde{\theta}) \equiv \rho(\widetilde{\theta}, G(I(\widetilde{\theta}))), \tag{41}$$

where $G(I(\widetilde{\theta}))$ is defined in (14), denote the fraction of a group assigned to the complex task if the firms use $\widetilde{\theta}$ as the assignment threshold. The affirmative action employment quota constraint (40) requires that $\hat{\rho}(\widetilde{\theta}_W) = \hat{\rho}(\widetilde{\theta}_B)$. In general $\hat{\rho}(\widetilde{\theta}_W) = \hat{\rho}(\widetilde{\theta}_B)$ does not necessarily imply $\widetilde{\theta}_W = \widetilde{\theta}_B$ because $\hat{\rho}(\cdot)$ may not be monotonic (as illustrated in the next section regarding "patronizing equilibrium"). How $\hat{\rho}(\cdot)$ varies with $\widetilde{\theta}$ depends on the interaction of two distinct effects. On the one hand, an increase in the threshold $\widetilde{\theta}$ makes it harder to be assigned to the complex task for a given fraction of qualified workers, thus leading to a decrease of $\hat{\rho}$; on the other hand, as $\widetilde{\theta}$ increases, the workers' skill investment incentives change, leading to changes in the fraction of qualified workers. The net effect is typically ambiguous. However, $\hat{\rho}(\cdot)$ must be decreasing over some part of the domain [0, 1] because $\hat{\rho}(0) = 1$ and $\hat{\rho}(1) = 0$. Thus a sufficient condition under which all equilibria under affirmative action entail homogeneous beliefs about the two groups is that $\hat{\rho}(\cdot)$ as defined in (41) is decreasing on [0, 1].

6.2.2 Patronizing equilibrium: an example

Coate and Loury (1993a) provided an example to demonstrate the possibility of patronizing equilibria under affirmative action. The idea is very simple: to comply with the affirmative action policy (assuming $\pi_B < \pi_W$ is unchanged by the policy for the moment), the standards for blacks must be lowered and the standards for whites must be raised to comply with the employment quota. Thus, it is now easier for blacks to be assigned to the good job (and harder for whites) irrespective of whether or not a particular worker invested in skills. Since the incentives to invest depend on the expected wage difference between skilled and unskilled workers, whether the above change will increase or decrease blacks' incentive to invest in skills depends on the particularities of the distributions f_q and f_u.

Consider the following example. Suppose that the skill investment cost c is uniform on [0, 1]. Assume the following test signal densities for qualified and unqualified workers, respectively:

$$f_q(\theta) = \begin{cases} \frac{1}{1-\theta_q} & \text{if } \theta \in [\theta_q, 1] \\ 0 & \text{otherwise,} \end{cases} \tag{42}$$

$$f_u(\theta) = \begin{cases} \frac{1}{\theta_u} & \text{if } \theta \in [0, \theta_u] \\ 0 & \text{otherwise,} \end{cases} \tag{43}$$

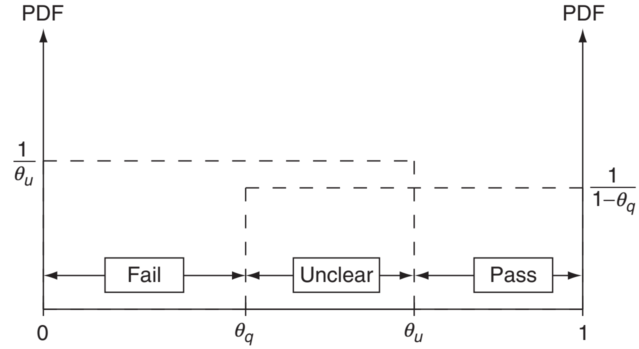

Figure 5 Signal distributions in Coate and Loury's (1993a) example of patronizing equilibrium.

where $\theta_u > \theta_q$. Figure 5 graphically illustrates these two distributions, which are equivalent to the case in which only three test results are possible. If $\theta > \theta_u$, then the signal is only possible if the worker is qualified, thus we call it a "pass" score; if $\theta < \theta_q$, then the signal is only possible if the worker is unqualified, thus we call it a "fail" score; if $\theta \in [\theta_q, \theta_u]$, then the signal is possibly from both a qualified and an unqualified worker, thus we call such a signal "unclear."

Equilibria without Affirmative Action. Let us first analyze the equilibrium of this example with no affirmative action. Clearly, the firm assigns workers with a "pass" score to the complex task and those with "fail" score to the simple task. Now we determine the optimal assignment decision regarding workers with "unclear" scores. It is clear from Figure 5 that the probability that a qualified worker gets an "unclear" score $\theta \in [\theta_q, \theta_u]$ is:

$$p_q = \frac{\theta_u - \theta_q}{1 - \theta_q}; \tag{44}$$

and for an unqualified worker is:

$$p_u = \frac{\theta_u - \theta_q}{\theta_u}. \tag{45}$$

Suppose that the prior that a worker is qualified is π. Then the posterior probability that a worker with an unclear score is qualified is, by Bayes' rule:

$$\xi(\pi) = \frac{\pi p_q}{\pi p_q + (1 - \pi) p_u}. \tag{46}$$

Hence, the employer will assign a worker with unclear scores to the complex task if and only if:

$$\xi(\pi) x_q - [1 - \xi(\pi)] x_u \geq 0,$$

or equivalently,

$$\pi \geq \hat{\pi} = \frac{p_u/p_q}{x_q/x_u + p_u/p_q}. \tag{47}$$

We say that a firm follows a *liberal* policy for group j if it assigns all group j workers with an unclear test score to the complex task, i.e., if $\widetilde{\theta} = \theta_q$; we say that a firm follows a *conservative* policy for group j if it assigns all group j workers with an unclear test score to the simple task, i.e., if $\widetilde{\theta} = \theta_u$.

In order for a liberal policy to be consistent with equilibrium, it must be the case that the skill investment incentives under the liberal policy will result in the fraction of qualified workers in the group to be larger than $\hat{\pi}$ defined in (47). Note that under a liberal policy, the benefit from skill investment is given by:

$$I(\theta_q) = \omega(1 - p_u)$$

because if the worker is skilled, he will be assigned with probability one to the complex task and if he is unskilled, the probability is p_u. Thus, the proportion of skilled workers in response to a liberal policy is:

$$\pi_l = I(\theta_q) = \omega(1 - p_u). \tag{48}$$

Thus the liberal policy is an equilibrium if $\pi_l > \hat{\pi}$.

Similarly, under a conservative policy, the benefit of skill investment is:

$$I(\theta_u) = \omega(1 - p_q).$$

Hence the proportion of skilled workers in response to a conservative policy is:

$$\pi_c = I(\theta_u) = \omega(1 - p_q). \tag{49}$$

Thus the conservative policy is an equilibrium if $\pi_c < \hat{\pi}$.

To summarize, in the absence of the affirmative action constraint, if $\pi_c < \hat{\pi} < \pi_l$, then the example admits multiple equilibria in that both the liberal policy and the conservative policy could be equilibria. Suppose that the blacks and the whites are coordinated on the conservative and the liberal equilibria, respectively; that is, $(\pi_B, \pi_W) = (\pi_c, \pi_l)$. Clearly, in this equilibrium, firms hold a negative stereotype toward blacks because $\pi_c < \pi_l$.

Equilibria with Affirmative Action. Suppose that the economy is in an equilibrium characterized by $(\pi_B, \pi_W) = (\pi_c, \pi_l)$ described above, and suppose that an affirmative action policy in the form of employment quota as described in Section 6.2.1 is imposed.[17]

[17] It can be verified that the sufficient condition for affirmative action to eliminate discriminatory equilibrium described in the previous section does not hold in this example.

Given that in the pre-affirmative action equilibrium $(\pi_B, \pi_W) = (\pi_c, \pi_l)$, there is a higher fraction of whites on the complex job. In order to comply with the affirmative action employment quota, the firm must either assign more blacks or assign fewer whites to the complex task. Which course of action is preferred will depend on the following calculations. Given $(\pi_B, \pi_W) = (\pi_c, \pi_l)$, if the firm assigns a black worker with a "fail" score to the complex task, it loses x_u unit of profits; however, if the firm assigns a white worker with an "unclear" score to the simple task (instead of the complex task as stipulated under the liberal policy), it loses:

$$\xi(\pi_l)x_q - [1 - \xi(\pi_l)]x_u,$$

where $\xi(\cdot)$ is defined in (46). Notice that if:

$$\lambda[\xi_l x_q - (1 - \xi_l)x_u] > (1 - \lambda)x_u,$$

then the firm would rather put all black workers with "fail" scores to the complex task than to switch white workers with "unclear" scores to the simple task in order to satisfy the employment quota.

Now consider the following assignment policies. For the whites, keep the original liberal policy; namely, assign all workers with "pass" or "unclear" scores to the complex task. Under this policy, the white workers' skill investment decisions in equilibrium will lead to $\pi_W = \pi_l$, same as before. For the black workers, the firms follow the following "*patronizing*" assignment policy: assign all black workers with "pass" or "unclear" scores to the complex task, *and* with probability $\alpha(\pi_B) \in (0, 1)$ assign blacks *with "fail" scores* to the complex task, where $\alpha(\pi_B)$ is chosen to satisfy the employment quota requirement:[18]

$$\alpha(\pi_B) = \frac{\pi_l - \pi_B}{1 - \pi_B}. \tag{50}$$

The firms are "*patronizing*" the blacks in this postulated assignment policy because they are assigning blacks who have "fail" scores to the complex task.

Now consider a black worker's best response if he anticipates being patronized with probability α. If he invests in skills, he will be assigned to the complex task with probability 1; if he does not invest, he will be assigned to the complex task with probability $p_u + (1 - p_u)\alpha$. Thus, the return from investing in skills for a black worker is:

$$\omega\{1 - [p_u + (1 - p_u)\alpha]\} = \omega(1 - \alpha)(1 - p_u) = (1 - \alpha)\pi_l$$

where the last equality follows from (48).

[18] That is, to satisfy.

$$\pi_l + (1 - \pi_l)p_u = \pi_B + (1 - \pi_B)[p_u + (1 - p_u)\alpha(\pi_B)].$$

Hence, any (π_B, π_l) where $\pi_l > 1/2$, can be sustained as an equilibrium under the affirmative action policy where firms follow a patronizing assignment policy $\alpha(\pi_B)$ for blacks and a liberal policy for whites if and only if $\pi_B \leq \pi_l$ and π_B satisfies:

$$\pi_B = [1 - \alpha(\pi_B)]\pi_l = \frac{(1-\pi_l)\pi_l}{1-\pi_B}. \tag{51}$$

Note that equation (51) admits two solutions for $\pi_B : \pi_B = \pi_l$ or $\pi_B = 1 - \pi_l$. In the first solution, color-blind equilibrium is reached and the employer is liberal toward both groups (at $\pi_B = \pi_l$, it can be seen from (50) that $\alpha(\pi_B) = 0$, thus there is no patronizing). In the second solution, the firms continue to view black workers as less productive in equilibrium and adopt a patronizing assignment policy on the blacks in order to fulfill the affirmative action employment quotas.

Dynamics. Coate and Loury (1993a) further argued that, under a plausible dynamics on the evolution of firms' beliefs about the fraction of blacks who invest in skills specified as system:

$$\begin{aligned} \pi_B^{t+1} &= [1-\alpha(\pi_B^t)]\pi_l \\ &= \frac{1-\pi_l}{1-\pi_B^t}\pi_l, \end{aligned}$$

with initial condition that $\pi_B^0 = \pi_c$, it can be shown using a simple phase diagram that $\pi_B^t \to 1 - \pi_l$ as $t \to \infty$. Thus in some sense, not only is the patronizing equilibrium possible, it could actually be a stable equilibrium outcome. Coate and Loury (1993b) studied the effect of affirmative action in a similar environment, but one where employers also hold prejudicial preferences against minorities. In that case, it is shown that a gradual policy in which representation targets are gradually increased might be more likely to eliminate disparities than radical policies demanding immediate proportional representation.

6.3 General equilibrium consequences of affirmative action

One weakness of Coate and Loury (1993a)'s model is that wages are not determined in a competitive labor market, but are fixed exogenously. Because affirmative action policies change the profitability of hiring workers from different groups, this is not an innocuous assumption. Moreover, workers from the discriminated group face a more favorable task assignment rule, but, conditional on the signal, receive the same wages as before, therefore affirmative action can only be a benefit to them.

Moro and Norman (2003a) study the effect of affirmative action policies in the general equilibrium setting analyzed in Section 4.1, where firms engaged in Bertrand competition for workers determine wages endogenously. Their analysis confirms the perverse incentive effects of government-mandated policies found by Coate and Loury (1993a).

Moreover, it finds perverse effects on equilibrium wages and proves that in some circumstances affirmative action may hurt its intended beneficiaries.

The affirmative action constraint is the same as that assumed in Section 6.2.1, that is, employers are forced to hire the same proportion of workers from both groups in the complex task (and, residually, in the simple task). Employers therefore solve the following problem (assuming for simplicity that groups have identical size):

$$\max_{\widetilde{\theta}_B,\ \widetilde{\theta}_W} y(C,S) = \max_{\widetilde{\theta}_B,\ \widetilde{\theta}_W} y\left(\sum_{j=B,W} \pi_j[1-F_q(\widetilde{\theta}_j)], \sum_{j=B,W}[\pi_j F_q(\widetilde{\theta}_j) + (1-\pi_j)F_u(\widetilde{\theta}_j)]\right)$$
$$\text{s.t. } \pi_B F_q(\widetilde{\theta}_B) + (1-\pi_B)F_u(\widetilde{\theta}_B) = \pi_W F_q(\widetilde{\theta}_W) + (1-\pi_W)F_u(\widetilde{\theta}_W).$$

Denote $\hat{\theta}_j(\boldsymbol{\pi}), j = B, W$ as the optimal group-specific cutoff rules that solve this problem for a given vector $\boldsymbol{\pi} = (\pi_B, \pi_W)$. Employers assign all workers with signal above such thresholds to the complex task, and all other workers to the simple task. Observe that from the constraint, it follows directly that if $\pi_B < \pi_W$ then $\hat{\theta}_B(\boldsymbol{\pi}) > \hat{\theta}_W(\boldsymbol{\pi})$. The direct (partial-equilibrium) effect of the policy on the task assignment rule is to force employers to lower the task assignment threshold for the discriminated group, and to raise the threshold for the dominant group. It can be proved that the equilibrium wages are:

$$\hat{w}_j(\theta;\pi) = \begin{cases} p(\hat{\theta}_j(\boldsymbol{\pi}),\pi_j)x_q(\hat{C},\hat{S}) & \text{for } \theta < \hat{\theta}_j(\boldsymbol{\pi}) \\ p(\theta,\pi_j)x_q(\hat{C},\hat{S}) & \text{for } \theta \geq \hat{\theta}_j(\boldsymbol{\pi}) \end{cases} \tag{52}$$

where $\hat{C}, \hat{S}$ are the optimal inputs of the production function computed from the optimization problem satisfying the affirmative action constraint, x_q and x_u are the marginal products of workers in the complex and simple task, and $p(\theta, \pi_j)$ is the probability that a worker with signal θ is qualified, given by (4). This result says that the wage is a continuous function of the signal, that workers in the complex task are paid exactly their marginal products, and that workers in the simple task are paid the wage of the marginal worker. In the simple task, workers are therefore paid above the marginal product if they belong to the dominant group and below their marginal product if they belong to the discriminated group. Figure 6 illustrates the equilibrium wages under the assumption $\pi_B < \pi_W$.

The proof of this result first argues that wages must be continuous, otherwise one employer could exploit the discontinuity and increase profit by offering a slightly higher wage to workers that are cheaper near the discontinuity, and zero to workers that are more expensive. Second, note that there is a difference between quantity of workers in the complex task and their labor input, because not all workers employed in the complex task are productive. If workers in the complex task were not paid their expected marginal product, then employers could generate a profitable deviation that exploits the difference between quantity of workers and quantity of effective

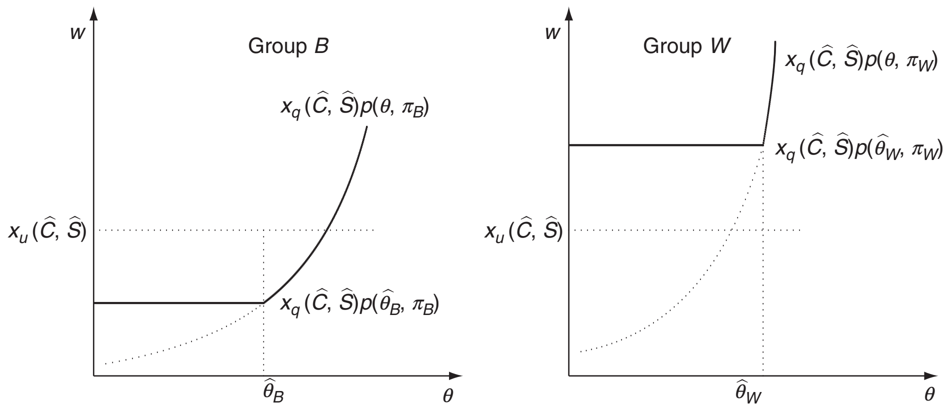

Figure 6 Equilibrium wage schedules under affirmative action in Moro and Norman (2003).

inputs.[19] However, because of continuity, this implies that workers in the simple task are paid above or below the marginal product depending on their group identity. It is not difficult to show from the first order condition of the task assignment problem that the average pay of all workers in the simple task (from both groups) is exactly the marginal product $x_u(\hat{C}, \hat{S})$.

Incentives to invest for group j are:

$$I_j(\boldsymbol{\pi}) = \int_\theta \hat{w}_j(\theta) f_q(\theta) d\theta - \int_\theta \hat{w}_j(\theta) f_u(\theta) d\theta, \; j = B, W \tag{53}$$

and the equilibria are characterized by the solution to the system of fixed-point equations $\pi_j = G(I_j(\boldsymbol{\pi}))$, $j = B, W$, where as usual G is the CDF of the cost of human capital investment. Any symmetric equilibrium of the model without the policy trivially satisfies the affirmative action constraint and therefore is also an equilibrium under affirmative action.

The full equilibrium effects of affirmative action are indeterminate. While it is possible that imposing affirmative action completely eliminates asymmetric equilibria, it is also possible for asymmetric equilibria to exist that satisfy the quota imposed by the policy for reasons similar to those illustrated by the patronizing equilibria derived in Section 6.2.2. A proof may be derived by construction by fixing fundamentals γ, f_q and f_u, and looking for a cost of investment distribution G that satisfies the equilibrium conditions under affirmative action. Note that if $\pi_B = 0$, and $0 < \pi_W < 1$, then from (52) and (53) it must be that $I_B(0, \pi_W) = 0 < I_W(0, \pi_W)$ (all group-B workers are offered zero wage, equivalent to their productivity in the complex task but some are employed in the complex task to satisfy the affirmative action constraint). But then since $I_j(\cdot)$ is

[19] The reader is invited to consult the proof in the original paper for details.

continuous and initially increasing near $\pi_B = 0$, one can find $\pi_B > 0$ such that $0 < \pi_B < \pi_W < 1$ and, at the same time, $0 < I_B(\pi_B, \pi_W) < I_W(\pi_B, \pi_W)$. Hence, one can find a strictly increasing CDF G such that $G(0) > 0$, $G(I_B(\pi_B, \pi_W)) = \pi_B$, and $G(I_W(\pi_B, \pi_W)) = \pi_W$ so that (π_B, π_W) is an equilibrium of the model.

In general, comparing outcomes with and without the policy is difficult because outcomes depend on the equilibrium selection. It is possible to show that the policy may have negative welfare effects for its intended beneficiaries. The negative direct effects on the discriminated group's wages are evident from Figure 6. The picture however hides the full equilibrium effects because factor ratios will change in equilibrium. Unless such factor ratios do not change significantly, expected earnings for group-B decrease. Note also from the figure that the direct effect of the policy is to increase incentives to invest for the discriminated group. This tends to moderate the negative wage effects, but unless this effect is significant, workers in the discriminated groups are made worse-off by the policy.

The wage determination in this model is specific to the modeling assumptions made regarding production and information technologies. In this simplified setting, a slightly more complex policy that combines affirmative action employment quota and racial equality of average wages in each task would be effective in inducing symmetric equilibria. It is not clear, however, whether such a policy would be easily implementable in a more complex environment.[20] Nevertheless, the model is useful to illustrate that affirmative action policies have non-trivial general equilibrium effects.

6.4 Affirmative action in a two-sector general equilibrium model

Fang and Norman (2006) derive similar, but more clear-cut, perverse results in a two-sector general equilibrium model motivated by the following puzzling observation from Malaysia. Since its independence from British colonial rule in 1957, Malaysia protected the Malays by entitling them to certain privileges including political power, while at the same time allowing the Chinese to pursue their economic objectives without interference. This relative racial harmony was rejected in 1970 when the so-called New Economic Policy was adopted, in which wide-ranging preferential policies favoring the Malays were introduced, most important of which is an effective mandate that only the Malays can access the relatively well-paid public sector jobs. However, despite the aggressive preferential policies favoring the Malays, the Malay did not achieve significant economic progress relative to the Chinese; if anything, the opposite seems to be true, that is, the new policy reversed the pre-1970 trend of the narrowing wage gaps between the Chinese and the Malays.

[20] Lundberg (1991), for example, describes how companies may use variables that are correlated with race to evade the imposition of policies that monitor the employment process, such as affirmative action. In that setting, it is shown that policies monitoring outcomes may be more effective in reducing inequality, at the cost of higher production losses from workers' misallocation.

Fang and Norman (2006) considered the following simple model. Consider an economy with two sectors, called respectively the *private* and the *public* sector. The private sector consists of two (or more) competitive firms, indexed by $i = 1, 2$. Firms are risk neutral and maximize expected profits, and are endowed with a technology that is complementary to workers' skills. A skilled worker can produce $x > 0$ units of output, and an unskilled one will, by normalization, produce 0.

The public sector offers a fixed-wage $g > 0$ to any worker who is hired, but there is rationing of public sector jobs: the probability of getting hired in the public sector if a worker applies is given by $\rho_j \in [0,1]$, where $j \in \{A, B\}$ is the worker's ethnic identity. In our analysis below, we treat ρ_j as the government's policy parameter. Government-mandated discriminatory policies are simply modeled by the assumption that $\rho_A \neq \rho_B$. Workers who apply for but are unsuccessful in obtaining public sector employment can return to and obtain a job in the private sector without waiting.

For each ethnic group $j \in \{A, B\}$, there is a continuum of workers with mass λ_j in the economy. Workers are heterogeneous in their costs, denoted by c, of acquiring the requisite skills for the operation of the firms' technology. The cost c is private information of the worker and is distributed according to a uniform [0, 1] distribution in the population of both groups. Workers are risk neutral and do not care directly about whether they work in the public or private sector. If a worker of cost type c receives wage w, her payoff is $w - c$ if she invests in skills, and w if she does not invest.

The events in this economy are timed as follows: In the first stage, each worker in group j with investment cost $c \in [0, 1]$ decides whether to invest in the skills. This binary decision is denoted by $s \in \{0, 1\}$ where $s = 0$ stands for no skill investment and $s = 1$ for skill acquisition. If a worker chooses $s = 1$, we say that she becomes *qualified* and hence she can produce β units of output in the private sector; otherwise she is *unqualified* and will produce 0. As in the other models surveyed in this section, skill acquisitions are *not* perfectly observed by the firms, but in the second stage the worker and the firms observe a noisy signal $\theta \in \{h, l\} \equiv \Theta$ about the worker's skill acquisition decision with the following distributions:

$$\Pr[\theta = h|s = 1] = \Pr[\theta = l|s = 0] = p < 1/2.$$

In the third stage, after observing the noisy signal θ, each worker decides whether to apply for the public sector job. If applying, she is accepted for employment in the public sector with probability ρ_j where j is her ethnic identity. If she was not employed in the public sector, she will, in the fourth stage, return to the private sector, where firms compete for her service by posting wage offers. After observing the wage offers, she decides which firm to work for, clearing the private sector labor market.

The key insight from Fang and Norman (2006) is that 'group j's incentives to invest in skills depend on the probability that they may receive the public sector employment ρ_j. To see this, suppose that at the end of the first stage, a proportion π_j of the group j

population is qualified. Then in the second stage, a total measure $p\pi_j + (1 - p)(1 - \pi_j)$ of workers receives signal h, among which a measure $p\pi_j$ is qualified and a measure $(1 - p)(1 - \pi_j)$ is unqualified. Similarly, a total measure $(1 - p)\pi_j + p(1 - \pi_j)$ of workers receives signal l, among which a measure $(1 - p)\pi_j$ is qualified and a measure $p(1 - \pi_j)$ is unqualified. Therefore, in the fourth stage, when a firm sees a group j worker with a signal θ, its posterior belief that this worker is qualified, denoted by $\Pr[s = 1 | \theta; \pi_j]$ where $\theta \in \{h, l\}$, is given by:

$$\Pr[s = 1|\theta = h; \pi_j] = \frac{p\pi_j}{p\pi_j + (1-p)(1-\pi_j)}$$
$$\Pr[s = 1|\theta = l; \pi_j] = \frac{(1-p)\pi_j}{(1-p)\pi_j + p(1-\pi_j)},$$

exactly as if there were no public sector. Hence, the equilibrium wage for group j workers with signal $\theta \in \{h, l\}$ when the proportion of qualified workers in group j is π_j, denoted by $w_\theta(\pi_j)$, is:

$$w_h(\pi_j) = \beta \Pr[s = 1|\theta = h; \pi_j] = \frac{\beta p\pi_j}{p\pi_j + (1-p)(1-\pi_j)}$$
$$w_l(\pi_j) = \beta \Pr[s = 1|\theta = l; \pi_j] = \frac{\beta(1-p)\pi_j}{(1-p)\pi_j + p(1-\pi_j)}.$$

Now we analyze the public sector job application decision in the third stage. A group j worker with signal θ applies to the public sector job if $w_\theta(\pi_j) < g$ and does not apply if $w_\theta(\pi_j) > g$ where g is the public sector wage. Defining $\hat{\pi}_\theta$ as the solution to $w_\theta(\hat{\pi}_\theta) = g$ for $\theta \in \{h, l\}$, i.e.,

$$\hat{\pi}_h = \frac{g(1-p)}{g(1-p) + p(\beta - g)}, \hat{\pi}_l = \frac{gp}{gp + (1-p)(\beta - g)}.$$

We can conclude that a group j worker with signal θ applies for a public sector job if and if $\pi_j \leq \hat{\pi}_\theta$.

A worker's incentive to acquire skills in the first stage comes from the subsequent expected wage differential between a qualified and an unqualified worker. With some algebra it can be shown that the incentive to invest in skills for group j workers, denoted by $I(\pi_j, \rho_j)$, is equal to the gain in expected wage from skill investment in the first stage relative to not invest, and is given by:

$$I(\pi_j, \rho_j) = \begin{cases} (2p-1)(1-\rho_j)[w_h(\pi_j) - w_l(\pi_j)] & \text{if } 0 \leq \pi < \hat{\pi}_h \\ (2p-1)\{(1-\rho_j)[w_h(\pi_j) - w_l(\pi_j)] + \rho_j[w_h(\pi_j) - g]\} & \text{if } \hat{\pi}_h \leq \pi < \hat{\pi}_l \\ (2p-1)[w_h(\pi_j) - w_l(\pi_j)] & \text{if } \hat{\pi}_l \leq \pi \leq 1. \end{cases} \quad (54)$$

Notice that the incentive to invest, $I(\pi_j, \rho_j)$, depends also on ρ_j, the probability of public sector employment for group j workers, which is the reason for a government-mandated preferential (or discriminatory) policy in the public sector to matter for the private sector labor market in our model. Indeed, a higher probability of public sector jobs will unambiguously decrease the investment incentives if $\pi < \hat{\pi}_l$ because:

$$\frac{\partial I(\pi_j, \rho_j)}{\partial \rho_j} = \begin{cases} -(2p-1)[w_h(\pi_j) - w_l(\pi_j)] < 0 & \text{if } \pi_j < \hat{\pi}_h \\ (2p-1)[w_l(\pi_j) - g] < 0 & \text{if } \hat{\pi}_h \leq \pi_j < \hat{\pi}_l \\ 0 & \text{otherwise.} \end{cases} \tag{55}$$

The intuition is simple: the public sector does not give any advantage to qualified workers over unqualified workers. As a result, a higher ρ_j always reduces the equilibrium level of π_j.

Now consider an economy where a minority ethnic group, say group A, is subject to government-mandated discrimination in the sense that $\rho_A = 0$; while the majority native group, group B, obtains public sector jobs with probability $\rho_B > 0$. Fang and Norman (2006) show that the discriminated group A, nevertheless, may be economically more successful than the preferred group B. Specifically, when the government marginally increases ρ_B from 0, there is a *direct effect* because now group B will have a higher degree of access to a higher paying public sector and they will less likely enter the private sector. If the public sector wage g is higher than the best private sector wage (i.e., $g > p\beta$), as assumed, this direct effect is a positive for group B. However, there is also a negative *indirect general equilibrium effect* because as ρ_B increases from 0, it also *reduces* the incentives of skill investment, which will in turn lower the expected wages in the private sector for group B. If g is not too high (i.e., $g < 4p\,(1-p)\beta$), then the expected wage of both qualified and unqualified group A workers are higher than those of respective group B workers if $\rho_A = 0$ and $\rho_B > 0$ is sufficiently small. Note that to satisfy the condition $p\beta < g < 4p\,(1-p)\,\beta$, the precision of the test signal p has to be less than 3/4. That the precision in the signal cannot be too high for the negative indirect effect to dominate should be intuitive: A beneficial net effect from being excluded from the public sector can only occur if the informational free riding problem in the private sector is severe enough; and the higher p, the less severe this problem is. It can also be shown that, under the same set of assumptions, not only group B workers have lower expected wages, but also group B workers of all skill investment cost types are economically worse off than their group A counterparts.

6.5 Role model effects of affirmative action

Advocates of affirmative action have often argued that larger representation of minorities in higher paying jobs and occupations can generate role models that can positively influence future generations of minorities in their investment decisions. Chung (2000) formalizes these arguments. Consider a group of individuals who differ in their costs of investment, which take on two possible values c_l or c_h with $c_l < c_h$. In the population, a fraction $\alpha \in (0, 1)$ is of type c_l. An individual's skill investment cost is her private information.

Table 2 Transition matrix of the probability of being hired to the complex job

$p_t \backslash p_{t+1}$	p_1	p_2
p_1	$1-\theta_{12}$	θ_{12}
p_2	θ_{21}	$1-\theta_{21}$

Each individual, upon learning her investment cost type c, makes a binary investment decision. The skill investment decision affects the probability that the individual will obtain a higher paying job. For simplicity, suppose that there are two kinds of jobs, a complex job that pays w and a simple job whose wage is normalized to 0. Suppose that $w > c_h > c_l > 0$.

If an individual invests in skills, then she will obtain the complex job with probability p that is drawn from a two-point distribution $\{p_1, p_2\}$ with $0 < p_1 < p_2 < 1$. Specifically, p follows a discrete-time Markov process as follows. The probability that $p = p_1$ in period 0 is equal to q_0, and q_0 is common knowledge among all individuals; the transition probability $\Pr(p_{t+1} = p_j | p_t = p_i)$ is given in Table 2 where both θ_{12} and θ_{21} lie in (0, 1/2).

Suppose that in each period, one individual makes an investment decision and then receives a job placement. All individuals observe the prior job placements of others, but do not observe their investment decisions.

To characterize the equilibrium investment decisions of the agents, the key is to characterize how the individuals' beliefs about the state of the labor market, whether p is equal to p_1 or p_2, evolve over time. The role model effect in this model refers to the phenomenon that a placement of a minority candidate in the high paying complex job will *increase* subsequent minorities' belief that the labor market condition for skilled workers is in state p_2, and as a result subsequent minorities' incentives to invest in skills increase.

Consider the first individual. Suppose that her belief about the state of the labor market at period 0 being $p = p_1$ is q_0. Assume for simplicity that the skill investment costs c_l and c_h are such that, at the belief that $p = p_1$ with probability q_0, an individual with investment cost c_l will invest in skills, but an individual with cost c_h will not. Moreover, consider a situation following a long history of individuals being placed on the simple job, and as a result the population's belief about the labor market being poor, i.e., $p = p_1$, is at a steady state $q^* \in (0, 1)$. That is, if another individual is observed to be placed on the simple job, the subsequent individual's belief about $p = p_1$ will stay at q^*.[21]

[21] Specifically, q^* solves the unique root in (0, 1) for the following quadratic equation:

$$\alpha(p_2 - p_1)(1 - \theta_{12} - \theta_{21})q^2 + [(\theta_{12} + \theta_{21})(1 - \alpha p_1) - \alpha(1 - \theta_{21})(p_2 - p_1)]q - \theta_{21}(1 - \alpha p_1) = 0.$$

The exact value of q^* can be easily derived from a steady state condition, and its expression is omitted here.

In the above situation, suppose that the n-th individual is the *very first* one who manages to land a complex job. Upon observing this, the $(n+1)$-st individual will now infer that the n-th individual had invested and thus must have had low skill investment cost. The posterior belief of the $(n+1)$-st individual that the state of the labor market in period n is $p = p_1$ is

$$q_n = \frac{[q^*(1-\theta_{12}) + (1-q^*)\theta_{21}]p_1}{[q^*(1-\theta_{12}) + (1-q^*)\theta_{21}]p_1 + [q^*\theta_{12} + (1-q^*)(1-\theta_{21})]p_2}.$$

It can be shown with some algebra that $q_n < q^*$, that is, upon the observation of a placement on the complex job, the future individuals' belief about the labor market improves. The n-th individual, upon being placed on the complex job, becomes a *role model* for future individuals. If c_h is not too high, this improvement in the belief may lead to those individuals with investment cost c_h to invest in skills as well. Thus a role model may lead to real changes in behavior among future generations. Chung (2000) also analyzed how long the role model effect may last.

However, if the role model effect is indeed an informational phenomenon, then once affirmative action is announced the beliefs of the disadvantaged group regarding the labor market should switch to $p = p_2$, thus there is no additional information about p being conveyed by preferential hiring in favor of the disadvantaged group. Hence, a standard role-model argument in favor of affirmative action is not supported when role-model effects are purely informational. Chung (2000) observes that only when the hiring of minorities have some payoff-relevant effect than anti-discriminatory policies can have a bite, for example when jobs require race-specific know-how, and there are so few minorities employed in positions requiring skills that the returns to such skills are uncertain among minorities.

6.6 Color sighted vs. Color blind affirmative action

6.6.1 Recent developments in the affirmative action policies related to college admission

Race-conscious affirmative action policies in college admission came under a lot of scrutiny ever since the landmark case of *Regents of the University of California vs. Bakke*, 438 U.S. 265 (1978) where the Supreme Court upheld diversity in higher education as a "compelling interest" and held that "race or ethnic background may be deemed a 'plus' in a particular applicant''s file" in university admissions, and at the same time ruled that quotas for underrepresented minorities violates the equal protection clause. In the 1996 case, *Hopwood vs. Texas* the Court banned any use of race in school admissions in Texas. To accommodate the ruling, the State of Texas passed a law guaranteeing entry to any state university of a student's choice if they finished in the top 10% of their graduating class.

Also in 1996, Proposition 209 was passed in California, which mandates that "the state shall not discriminate against, or grant preferential treatment to, any individual or group on the basis of race, sex, color, ethnicity, or national origin in the operation of public employment, public education, or public contracting."[22] Proposition 209 essentially prohibits public colleges and universities in California from using race in any admission or financial aid decision. From 2001, the top 4% of high school seniors are guaranteed admission to any University of California campus under California's Eligibility in Local Context plan. In 1998, Washington state voters overwhelmingly passed Initiative 200, which is almost identical to California's Proposition 209. Florida passed its Talented 20 Plan, which guaranteed Florida high school students who graduate in the top 20% of their class admissions to any of the eleven public universities within the Florida State University System.

Two 2003 Supreme Court cases on affirmative action in admissions are related to the University of Michigan. In *Grutter vs. Bollinger*, the Supreme Court upheld the affirmative action admissions policy of the University of Michigan Law School. The Court's majority ruling, authored by Justice Sandra Day O'Connor, held that the United States Constitution "does not prohibit the law school's narrowly tailored use of race in admissions decisions to further a compelling interest in obtaining the educational benefits that flow from a diverse student body." In *Gratz vs. Bollinger*, on the other hand, the Supreme Court ruled that "the University [of Michigan]'s policy, which automatically distributes 20 points, or one-fifth of the points needed to guarantee admission, to every single 'underrepresented minority' applicant solely because of race, is not narrowly tailored to achieve educational diversity." On the one hand, the court affirmed that the use of race in admission decision is not unconstitutional, but at the same time, in the second case, the court specified that any automatic use of race in the computation of a scoring system used in determining admissions violate the constitution.

6.6.2 Color sighted vs. Color blind affirmative action with exogenous skills

Chan and Eyster (2003) studied the effect of color-blind affirmative action policies on the quality of admitted students when colleges have preferences for diversity.

Applicants. Consider a college who must admit a fraction C of applicants. The applicants belong to two groups, black (B) and white (W), with measure λ_B and λ_W respectively such that $\lambda_B + \lambda_W = 1$. Suppose that the test scores of the applicants (also exchangeably the quality of the applicants), denoted by $t \in [\underline{t}, \overline{t}]$, in group $j \in \{B, W\}$ is drawn from distributions $f_j(\cdot)$, such that $\int_{\underline{t}}^{\overline{t}} f_j(t)dt = 1$. Suppose that black applicants tend to have lower test scores than white applicants.[23] Specifically, assume that the distributions $f_W(\cdot)$ and $f_B(\cdot)$ satisfy the following strict monotone likelihood ratio property:

Assumption 2. f_W (t) / f_B(t) is continuously differentiable and strictly increasing in t for $t \in (\underline{t}, \overline{t})$.

[22] See http://vote96.sos.ca.gov/Vote96/html/BP/209text.htm

[23] See Fryer and Loury (2008), discussed below, for a model that links the distributions of test scores to ex ante investment efforts.

A key implication of this assumption is that higher test scores are more likely coming from white applicants.

Admissions. The admission office observes the applicants' test scores and their group identities, and makes admission decisions subject to the constraint that the fraction of applicants admitted must equal the capacity of the university C. Formally, an admission rule is (r_B, r_W), where $r_j(t) : [\underline{t}, \bar{t}] \rightarrow [0,1], j \in \{B, W\}$ is the probability that a group j member with test score t is accepted, such that $t_j(\cdot)$ is weakly increasing in t. The *admissible* admission rules depend on whether affirmative action is allowed. If it is allowed, then $r_j\ (t)$ can depend on j; if it is not allowed, then $r_B\ (t) = r_W\ (t)$ for all $t \in [\underline{t}, \bar{t}]$. For simplicity, let $N_j(r) = \lambda_j \int_{\underline{t}}^{\bar{t}} r_j(t) f_j(t) dt$ denote the number of group j applicants admitted under rule r.

The admission office's preference is postulated as a weighted average of the total test scores of admitted students and racial diversity. Specifically,

$$U(r) = \sum_{j \in \{B,W\}} \lambda_j \int_{\underline{t}}^{\bar{t}} t r_j(t) f_j(t) dt - \alpha \left| \lambda_B - \frac{N_B(r)}{C} \right| \tag{56}$$

where $\alpha > 0$ captures the admission office's taste for diversity; in particular, the university desires to achieve a racial composition in the student body that is identical to the racial composition of the applicant pool. Note that under (56) the admission office wants to achieve racial diversity whether or not the admission rules have to be -color-blind or are allowed to be color-sighted.

The admission office chooses $\langle r_B\ (t), r_W\ (t)\rangle$ among admissible set of admission rules to maximize (56) subject to the constraint that the capacity is reached, i.e.,

$$\sum_{j \in \{B,W\}} \lambda_j \int_{\underline{t}}^{\bar{t}} r_j(t) f_j(t) dt = C. \tag{57}$$

It is clear that restricting the admission office to color-blind admission rules will necessarily lower its attainable payoff; the goal of the analysis is to show how such color-blindness restriction affects the constrained optimal admission rules, and how it affects the test scores of admitted students, i.e., the first term in (56).

Color-Sighted Affirmative Action. When color sighted affirmative action is admissible, the admission office sets a cutoff rule for each group and admits any applicants scoring above her group's cutoff. Let (t_B^*, t_W^*) denote the admission test score threshold for black and white applicants respectively. If we ignore the absolute-value sign in the objective function (56), the admission office solves:

$$\max_{\{t_B, t_W\}} \lambda_B \int_{t_B}^{\bar{t}} \left(t + \frac{\alpha}{C}\right) f_B(t) dt + \lambda_W \int_{t_W}^{\bar{t}} t f_W(t) dt$$

subject to the capacity constraint. If the solution to the above modified problem has the minority group underrepresented, then ignoring the absolute-value sign is not consequential and the solution also solves the original problem. The first order conditions for the above modified problem with respect to t_B and t_W imply that:

$$t_B + \frac{\alpha}{C} = t_W.$$

If under such thresholds (t_B, t_W), minorities are indeed underrepresented, then we have a solution. If minorities are overrepresented, then the solution to the original problem will be thresholds that exactly achieve proportional representation. Thus, given Assumption 2, the optimal color sighted admission rule is a cutoff rule (t_B^*, t_W^*) such that $0 \leq t_W^* - t_B^* \leq \alpha/C$. Blacks are weakly underrepresented.

Color Blind Affirmative Action. A ban on color-sighted affirmative action would require that the same admission rule be used for both groups. Thus, the strict monotone likelihood ratio property would necessarily imply that the minority group will be under-represented among the admitted students as long as the admission rule is increasing in t. Hence the term $\alpha|\lambda_B - \frac{N_B(r)}{C}|$ in the admission office's objective function is simply $\alpha(\lambda_B - \frac{N_B(r)}{C})$. Dropping the constant $\alpha\lambda_B$ and using the fact that $N_B(r) = \lambda_B \int_{\underline{t}}^{\bar{t}} r(t) f_B(t)dt$, we can rewrite the admission office's problem as:

$$\begin{aligned} &\max_{r(\cdot)} U(r) = \int_{\underline{t}}^{\bar{t}} r(t)\gamma(t)[\lambda_B f_B(t) + \lambda_W f_W(t)]dt \\ &\text{s.t.} \sum_{j\in\{B,W\}} \lambda_j \int_{\underline{t}}^{\bar{t}} r(t) f_j(t)dt = C \end{aligned} \tag{58}$$

where,

$$\gamma(t) \equiv t + \frac{\alpha}{C}\frac{\lambda_B f_B(t)}{\lambda_B f_B(t) + \lambda_W f_W(t)} \tag{59}$$

The function γ defined above represents the increase in the admission office's utility from admitting a candidate with test score t. The first term is its utility from the test score itself, and the second term reflects its taste for diversity. Note that the likelihood that a test score of t is coming from a black applicant is given by the likelihood ratio $\lambda_B f_B(t) / [\lambda_B f_B(t) + \lambda_W f_W(t)]$.

The admission office obviously would like to fill its class with applicants with the highest value of γ. When γ is everywhere increasing in t, it can simply use a threshold rule. The problem is that γ might not be monotonic in t. To see this, note that the monotone likelihood ratio property implies that the second term in the expression $\gamma(\cdot)$ in (59) is strictly decreasing in t, but, in general, nonlinearly, which implies that γ might not be monotonic.

If $\gamma(\cdot)$ is not everywhere increasing in t, the admission office is not able to admit its favorite applicants without violating the constraint that $r(\cdot)$ must be increasing in t.

Chan and Eyster (2003) provides a useful characterization for the optimal color blind admission rule in this case. To describe their characterization, define $\Gamma\ (t_1,\ t_2)$ as the average value of γ over the interval $(t_1,\ t_2)$:

$$\Gamma(t_1, t_2) \equiv \begin{cases} \frac{\int_{t_1}^{t_2} \gamma(t)[\lambda_B f_B(t) + \lambda_W f_W(t)]dt}{\int_{t_1}^{t_2} [\lambda_B f_B(t) + \lambda_W f_W(t)]dt} & \text{for } t_1 < t_2 \\ \gamma(t_1) & \text{for } t_1 = t_2 \end{cases}.$$

The curves $\gamma(\cdot)$ and $\Gamma(\cdot, \bar{t})$ as a function of t are illustrated in Figure 7. In Figure 7, γ attains its maximum at t_a, but since r must be increasing in t, the admission office cannot admit applicants with test score t_a without also admitting students with higher test scores, even though as shown in the figure, those with higher test scores have lower values of γ. The optimal colorblind admission rule turns out to involve randomization and the optimal random rule depends on $\Gamma(\cdot, \bar{t})$. In Figure 7, $\Gamma(\cdot, \bar{t})$ attains the global maximum at t_m. Thus, the admission office prefers a randomly drawn applicant scoring above t_m to a randomly drawn applicant scoring above other t. If the capacity C is sufficiently small, the admission office will randomly admit applicants with test scores in the interval $[t_m,\ t]$ with a constant probability chosen to fill the capacity. If the capacity is sufficiently large, the admission office will admit all applicants with test scores above t_m with probability 1 and then admit applicants scoring below t_m in descending order of the test score. To summarize, if $\gamma(\cdot)$ as defined in (58) is not everywhere increasing in t, the optimal color blind admission rule must involve randomization for some values of capacity C.

Under random admission rules, applicants with higher test scores are not admitted with probability 1 at the same time that those with lower test scores are admitted with positive probability, the allocation of the seats are thus not efficient in terms of student quality. For any random colorblind admission rule r, one can construct a color sighted threshold admission rule $(t_B,\ t_W)$ that achieves the same diversity as that under r, but yields higher quality.

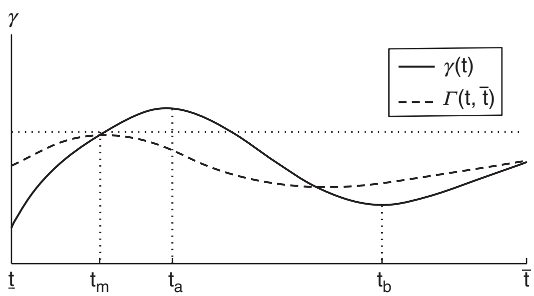

Figure 7 Admission office's preferences over test scores under color-blind admission policy (Figure 1 in Chan and Eyster 2003).

A general equilibrium framework. A similar analysis of the effect of banning Affirmative Action in college admissions, but with colleges competing for students, can be found in Epple, Romano, and Sieg (2008).[24] In their model, colleges care about the academic qualifications of their students and about income as well as racial diversity. Ability and income are correlated with race. Vertically differentiated colleges compete for desirable students using financial aid and admission policies. They show that because of affirmative action minority students pay lower tuition and attend higher-quality schools. The paper characterizes the effects of a ban on affirmative action. A version of the model calibrated to U.S. data shows that a ban of affirmative action leads to a substantial decline of minority students in the top-tier colleges. In an empirical analysis, Arcidiacono (2005) also finds that removing advantages for minorities in admission policies substantially decreases the number of minority students at top tier schools.

6.6.3 Color sighted vs. Colorblind affirmative action with endogenous skills

The analysis of affirmative action in Coate and Loury (1993a) assumed that quotas are to be imposed in the hiring stage. In practice, policymakers who are interested in improving the welfare of the disadvantaged group could potentially intervene in several different stages. For example, in the context of Coate and Loury's model, policymakers could potentially intervene by subsidizing the skill investment of workers from the disadvantaged group. Fryer and Loury (2008) extends the Chan and Eyster (2003) model to add an ex-ante skill investment stage to shed some light on the following question: "Where in the economic life-cycle should preferential treatment be most emphasized; before or after productivities have been determined?"

Recall that in Chan and Eyster (2003)'s model, the test score distribution for group j applicants are assumed to differ by group exogenously. Fryer and Loury (2008) endogenize the differences in $f_j(t)$ by assuming that groups differ in the distribution of investment costs, and that the test score distributions $f_j(t)$ are related to the investment decisions.

Specifically, let $G_j(c)$ be the cumulative distribution of skill investment cost in group j, and let $G(c) \equiv \sum_{j=\{B,W\}} \lambda_j G_j(c)$ be the effort cost distribution in the entire population, with $g_j(\cdot)$ and $g(\cdot)$ as their respective densities.

Denote an agent's skill investment decision as $e \in \{0, 1\}$. Suppose that the distribution of productivity v, analogous to the test score t in Chan and Eyster (2003), for an agent depends on e, with $H_e(v)$ and $h_e(v)$ as the CDF and PDF of v if the investment decision is e. If the fraction of individuals in group j who invested in skills is π_j, then the distribution of test scores in group j, again denoted by $F_j(v)$, with f being the corresponding density, will be:

$$F_j(v) \equiv F(v; \pi_j) = \pi_j H_1(v) + (1 - \pi_j) H_0(v).$$

[24] See also Epple, Romano and Sieg (2002).

Let $F^{-1}(z;\pi)$ for $z \in [0,1]$ denote the productivity level at the z-th quantile of the distribution $F(v;\pi)$. Suppose that there is a total measure $C<1$ of available "slots" that will allow an individual with productivity v to produce v units of output.

Laissez-faire Equilibrium. Fryer and Loury (2008) first analyzed the equilibrium allocation of the productive "slots" and the investment decisions under *laissez-faire*. Let π^m be the fraction of the population choosing $e=1$ in equilibrium and let p^m be the equilibrium price for a "slot." Clearly,

$$p^m = F^{-1}(1-C;\pi^m). \tag{60}$$

Given p^m, the ex-ante expected gross return from skill investment is:

$$\int_{p^m}^{\infty}(v-p^m)d\Delta H(v) = \int_{p^m}^{\infty}\Delta H(v)dv \tag{61}$$

where $\Delta H(v) = H_1(v) - H_0(v) \geq 0$. Since agents will invest in skills if and only if the expected gross return from skill investment exceeds the investment cost c, we have the following equilibrium condition:

$$\pi^m = G\left(\int_{p^m}^{\infty}\Delta H(v)dv\right). \tag{62}$$

The *laissez-faire* equilibrium (π^m, p^m) is thus characterized by equations (60) and (62). Note that after substituting the expression of p^m in (60) into (62), and taking G^{-1} on both sides, we have that the *laissez-faire* equilibrium of π^m must satisfy:

$$G^{-1}(\pi^m) = \int_{F^{-1}(1-C;\pi^m)}^{\infty}\Delta H(v)dv, \tag{63}$$

It can be formally shown that the *laissez-faire* equilibrium (π^m, p^m) characterized above is socially efficient. To see this, write an allocation as $\langle e_j(c), \alpha_j(v)\rangle$ where $e_j(c) \in \{0,1\}$, $\alpha_j(v) \in [0,1]$ are respectively the effort and slot assignment probability for each type of agent at the two stages. Let $\pi_j \equiv \int_0^{\infty} e_j(c)dG_j(c)$ be the fraction of group j population that invest in skills under effort rule $e_j(c)$. An allocation $\langle e_j(c), \alpha_j(v)\rangle$, $j \in \{B, W\}$, is feasible if:

$$\sum_{j\in\{B,W\}}\lambda_j\int\alpha_j(v)dF(v;\pi_j) \leq C. \tag{64}$$

An allocation is socially efficient if it maximizes the net social surplus:

$$\sum_{j\in\{B,W\}}\lambda_j\left[\int v\alpha_j(v)dF(v;\pi_j) - \int_0^{\infty}ce_j(c)dG_j(c)\right] \tag{65}$$

subject to the feasibility constraint (64).

We can rewrite the above efficiency problem as follows. Suppose that the fraction of agents investing in skills in some allocation is $\pi \in [0, 1]$, i.e., $\pi = \sum_{\lambda \in \{B,W\}} \int_0^\infty e_j(c) dG_j(c)$. Efficiency would require that the slots are only allocated to those in the top C quantile of the productivity distribution, thus the aggregate production for any given π in an efficient slot allocation rule must be:

$$Q(\pi) = \int_{1-C}^{1} F^{-1}(z;\pi)dz. \tag{66}$$

To achieve a fraction π of population investing, the efficient investment rule $e_j(c)$, $j \in \{B, W\}$, must be that only those in the lowest π-quantile in the effort cost distribution $G(\cdot)$ invest in skills. Thus the least aggregate effort costs to achieve π is:

$$C(\pi) = \int_0^{\pi} G^{-1}(z)dz. \tag{67}$$

Thus the socially efficient π is characterized by the first order condition $Q'(\pi) = C'(\pi)$, which yields:

$$G^{-1}(\pi^*) = \int_{1-C}^{1} \frac{\partial F^{-1}(z;\pi)}{\partial \pi} dz = \int_{F^{-1}(1-C;\pi^*)}^{\infty} \Delta H(v)dv. \tag{68}$$

The characterization for the socially efficient level of π^* is identical to that of the *laissez-faire* equilibrium of π^m provided in (63), thus $\pi^* = \pi^m$. Since it is also obvious that the slot assignment rule under the *laissez-faire* equilibrium allocation is exactly the same as the efficient assignment rule for a given π, we conclude that the *laissez-faire* equilibrium is efficient.

Let ρ_j^* be the faction of group j agents who acquires slots under the *laissez-faire* equilibrium. Under the plausible assumption that $g_B(c)/g_W(c)$ is strictly increasing in c, which, among other things, implies that $G_B(c)$ first order stochastically dominates $G_W(c)$, then the *laissez-faire* equilibrium will have a smaller fraction of the group B agents assigned with slots.

Let us suppose that a regulator aims to raise the fraction of group B agents with slots to a target level $\rho_B \in (\rho^*{}_B, C]$. Moreover, suppose that the regulator's affirmative action policy tools are limited to $(\sigma_W, \sigma_B, \tau_W, \tau_B)$ where σ_j is the regulator's transfers to group j agents who invest in skills and τ_j is a transfer to group j agents who hold slots. Fryer and Loury (2008) interpret σ_j as intervention at the *ex-ante investment margin*, and τ_j as intervention at the *ex post assignment margin*. It is easy to see that we can without loss of generality set either τ_W or τ_B to zero, because a universal transfer to all slot holders will just be capitalized into the slot price. Let us set $\tau_W = 0$.

Color-Sighted Intervention. First consider the case of color-sighted affirmative action, which simply means that (σ_j, τ_j) can differ by group identity j. Fix a policy

$(\sigma_W, \sigma_B, \tau_B)$, let π_j be the fraction of group j agents who invest in skills, and let p be the equilibrium slot price. We know that only group B agents with v above $p - \tau_B$ will obtain a slot. Thus to achieve the policy goal ρ_B, we must have

$$1 - F(p - \tau_B; \pi_B) = \rho_B,$$

that is,

$$p - \tau_B = F^{-1}(1 - \rho_B; \pi_B). \tag{69}$$

From the slot clearing condition, $\lambda_W \rho_W + \lambda_B \rho_B = C$, we can solve for ρ_W for any policy goal ρ_B, i.e., $\rho_W = (C - \lambda_B \rho_B)/\lambda_W$. The equilibrium slot price p must satisfy:

$$1 - F(p; \pi_W) = \rho_W,$$

or equivalently;

$$p = F^{-1}(1 - \rho_W; \pi_W). \tag{70}$$

A group j agent will invest in skills if his investment cost c, minus the transfer σ_j, is less than the expected benefit from investing. This gives us:

$$\pi_W = G_W\left(\sigma_W + \int_p^{\infty} \Delta H(v)dv\right) \tag{71}$$

$$\pi_B = G_B\left(\sigma_B + \int_{p-\tau_B}^{\infty} \Delta H(v)dv\right) \tag{72}$$

For a given pair (π_W, π_B), Equations (69)–(72) uniquely determine the policy parameters $(\sigma_W, \sigma_B, \tau_B)$ and the equilibrium slot price p for whites that will implement the affirmative action target $\rho_B \in (\rho^*{}_B, C]$. What remains to be determined is the constrained efficient levels of (π^s_W, π^s_B) which maximize the social surplus from implementing the policy objective (ρ_W, ρ_B), given by:[25]

$$\sum_{j \in \{B,W\}} \lambda_j \left[\int_{1-\rho_j}^{1} F^{-1}(z; \pi^s_j)dz - \int_0^{\pi^s_j} G_j^{-1}(z)dz\right]. \tag{73}$$

Problem (73) is separable by group. Thus, the first order condition for the constrained efficient levels of (π^s_W, π^s_B) is analogous to (68), except that now it is group specific, namely, for $j = B, W$,

[25] (73) is derived analogous to (66) and (67). Note that the transfers and subsidies $(\sigma_W, \sigma_B, \tau_B)$ do not factor into the calculation for social surplus.

$$G_j^{-1}(\pi_j^{s*}) = \int_{F^{-1}(1-\rho_j;\pi_j^{s*})}^{\infty} \Delta H(v)dv. \tag{74}$$

Combining the characterization of (π_W^{s*}, π_B^{s*}) provided in (74) with the (69)–(72), we immediately have the following result: given an affirmative action target $\rho_B \in (\rho^*{}_B, C]$, the efficient color sighted affirmative action policy is:

$$\sigma_W = \sigma_B = 0, \tau_B = F^{-1}(1-\rho_W; \pi_W^{s*}) - F^{-1}(1-\rho_B; \pi_B^{s*}),$$

where $\rho_W = (C - \lambda_B\, \rho_B)/\lambda_W$, and (π_W^{s*}, π_B^{s*}) satisfy (74).

In other words, when the affirmative action policies can be conditioned on group identity, the regulator will not use *explicit* skill subsidies to promote the access of a disadvantaged group to scarce positions. Of course, by favoring disadvantaged group at the slot assignment stage, skill investment is still *implicitly* subsidized for the disadvantaged. To spell out the intuition for the result, it is useful to note that, due to the noise in the productivity following skill investment, because productivity *vs.* conditional on investment is distributed as H_1 (v), subsidy on the *ex-ante* skill investment will lead to *leakage* in the sense that some black agents may decide to invest in skills as a result of skill subsidy, but may end up with low productivity and be assigned a slot. An *ex post* subsidy on the slot price for the blacks is a more *targeted* policy.

Color Blind Intervention. Now consider the case where policies cannot condition on color, that is, $\sigma_W = \sigma_B = \sigma^c$ and $\tau_W = \tau_B = \tau$. As we discussed earlier, if $\tau > 0$, but the price of slots are allowed to be set in equilibrium, the slot price subsidy τ will be reflected in a higher slot price. Thus in fact, the regulator may as well set $\tau = 0$, but instead *impose a cap* p^c *for the slot price*. The idea of implementing affirmative action using color blind policy instruments is similar to that detailed in Chan and Eyster (2003): imposing a lower threshold (i.e., a cap on the slot price) and employing randomization. If there are more blacks at the assignment margin p^m identified for the *laissez-faire* equilibrium, the affirmative action goal ρ_B may be achieved because lowering the margin and randomizing the slot assignment for those above the margin favors the blacks.

Let (σ^c, p^c) be the colorblind policy. Suppose that the fraction of individuals who invest in skills under such a policy is π^c in the population and π_j^c within group j. Given the price cap p^c, the total measure of individuals whose productivity v (and thus willingness to pay for a slot) is above p^c is given by $1 - F(p^c; \pi^c)$. Thus the random rationing probability, denoted by α^c, is given by:

$$\alpha^c = \frac{C}{1 - F(p^c; \pi^c)} < 1. \tag{75}$$

The gross returns from investing in skills when slots are rationed is given by $\sigma + \alpha^c \int_{p^c}^{\infty} \Delta H(v)dv$. Thus, the fractions of individuals who invest in skills are:

$$\pi^c = G\left(\sigma + \alpha^c \int_{p^c}^{\infty} \Delta H(v) dv\right), \tag{76}$$

$$\pi_j^c = G_j\left(\sigma + \alpha^c \int_{p^c}^{\infty} \Delta H(v) dv\right) = G_j(G^{-1}(\pi^c)) \text{ for } j = B, W. \tag{77}$$

In equilibrium, the proportion of blacks assigned with a slot is given by $\alpha^c[1 - F(p^c; \pi_B^c)]$. To satisfy the affirmative action target ρ_B, it must be the case that:

$$\rho_B = \alpha^c[1 - F(p^c; \pi_B^c)]. \tag{78}$$

Substituting the expression of α^c from (75) into (78), the affirmative action target constraint can be rewritten as:

$$\rho_B = \frac{C[1 - F(p^c; \pi_B^c)]}{1 - F(p^c; \pi^c)} = \frac{C[1 - F(p^c; G_B(G^{-1}(\pi^c)))]}{1 - F(p^c; \pi^c)}, \tag{79}$$

where the second equality follows from substituting (77) for π_B^c. It can be shown that, for a fixed π^c (and thus fixed π_B^c as well due to (77)), the right hand side is strictly decreasing in p^c. Thus for any target ρ_B, there exists a unique p^c to achieve the target and the price cap p^c is lower, the more aggressive the target ρ_B is.

Because (76) tells us that the skill subsidy σ^c is uniquely determined by (π^c, p^c), we can recast the regulator's problem as choosing (π^c, p^c) to maximize the social surplus given by:

$$\frac{C}{1 - F(p^c; \pi^c)} \int_{p^c}^{\infty} v dF(v; \pi^c) - \int_0^{\pi} G^{-1}(z) dz \tag{80}$$

subject to the affirmative action target constraint (79). Let (π^{c*}, p^{c*}) be the solution to the above problem. From the first order condition to problem (80), Fryer and Loury (2008) showed that σ^{c*} corresponding to (π^{c*}, p^{c*}), which can be derived from (76) as:

$$\sigma^{c*} = G^{-1}(\pi^{c*}) - \frac{C}{1 - F(p^{c*}; \pi^{c*})} \int_{p^{c*}}^{\infty} \Delta H(v) dv$$

is positive if and only if:

$$\frac{f(p^{c*}, G_B(G^{-1}(\pi^{c*})))}{f(p^{c*}, \pi^{c*})} < \frac{g_B(G^{-1}(\pi^{c*}))}{g(G^{-1}(\pi^{c*}))}. \tag{81}$$

Note that the left-hand side term, if multiplied by λ_B, is the relative fraction of blacks among agents on the *ex post* assignment margin p^c; and the right-hand side term, if multiplied by λ_B, is the relative fraction of blacks on the *ex-ante* skill investment margin with $c = G^{-1}(\pi)$. Thus, we have the following result: Given an affirmative action target $\rho_B \in (\rho^*{}_B, C]$, and let (π^{c*}, p^{c*}) solve problem (80), then the efficient color blind affirmative action policy will involve strictly positive skill investment subsidy $\sigma^{c*} > 0$ if (81) holds at (π^{c*}, p^{c*}).

6.7 Additional issues related to affirmative action

Besides the theoretical examinations of the effects of affirmative action on incentives and welfare, a recent literature asks whether affirmative action policies in college and professional school admissions may have led to mismatch that could inadvertently hurt, rather than, help, the intended beneficiaries. This so-called mismatch literature examines how some measured outcomes, such as GPA, wages, or bar passage rate, etc., for minorities are affected by affirmative action admission policies.[26] A recent paper by Arcidiacono, Aucejo, Fang, and Spenner (2009) takes a new viewpoint by asking why minority students would be willing to enroll themselves at schools where they cannot succeed, as stipulated by the mismatch hypothesis. They show that a *necessary condition* for mismatch to occur once we take into account the minority students' rational enrollment decisions is that the selective university has private information about the treatment effect of the students, and provide tests for the necessary condition. They implement the test using data from the Campus Life and Learning (CLL) project at Duke University. Evidence shows that Duke does possess private information that is a statistically significant predictor of the students' post-enrollment academic performance. Further, this private information is shown to affect subjective measures of students' satisfaction as well as their persistence in more difficult majors. They also propose strategies to evaluate more conclusively, whether the presence of Duke's private information has generated mismatch.

In the class of models where discriminatory outcomes arise because of multiple equilibria and coordination failure, as reviewed in Sections 3 and 4, affirmative action can be interpreted as an attempt to eliminate the Pareto dominated equilibrium where the disadvantaged group coordinates on. One of the problems, as illustrated by the patronizing equilibrium identified by Coate and Loury (1993a) and described in Section 6.2.2, is that affirmative action policies may lead to new equilibrium with inequality. In an interesting paper, Chung (1999) interprets the affirmative action problem as an *implementation* problem and ask whether more elaborate affirmative action policies can be identified that will eliminate the Pareto dominated equilibrium without generating any new undesirable equilibria. Chung (1999) shows that in a Coate and Loury model, a class of policies that combine unemployment insurance and employment

[26] See Loury and Garman (1995), Sanders 2004, Ayres and Brooks 2005, Ho (2005), Chambers et al. (2005), Barnes (2007), and Rothstein and Yoon (2008).

subsidy (insurance-cum-subsidy) can eliminate the bad equilibrium without generating any new undesirable equilibria. The insurance-cum-subsidy policy can be interpreted as follows: each worker from a certain group is offered an option to buy an unemployment insurance package at the time he makes his human capital investment. The insurance is unattractive to any worker unless the probability of being unemployed is sufficiently high; enough workers buying this insurance will trigger a group-wide employment subsidy. A policy like this does not lead to undesirable patronizing equilibrium because the employment subsidies appear only if workers believe the employers are too reluctant to hire them.

Abdulkadiroglu (2005) studies the effect of affirmative action in college admission from the perspective of matching theory. He interprets the college admissions problem as a many-to-one two-sided matching problem with a finite set of students and a finite set of colleges. Each college has a finite capacity to enroll students. The preference relation of each student over colleges is a linear order of colleges, where as the preference relation of each college over sets of students is a linear order of the set of students. He examines the conditions for the existence of stable mechanisms that make truthful revelation of student preferences a dominant strategy with and without affirmative action quotas.

Fu (2006) studies the effect of affirmative action using insights from all-pay auctions. He considers a situation where two students, one majority and one minority, are competing for one college seat. The college wants to maximize test scores, which depends only on the students' efforts. Suppose that the benefit from attending the college is higher for the majority student than for the minority student. The two students compete for the college seat by choosing effort levels. Fu (2006) shows that this problem is analogous to a asymmetric complete information all-pay auction problem where the college can be thought of as the "seller," and the two students the "bidders," the test scores (or the efforts) are the "bids," and the students' benefit from attending the college "values of the object to the bidders." He then uses insights from asymmetric all-pay auctions to show that to maximize the test scores; the college actually should adopt an admission rule that favors the minority students to offset his disadvantage in value from attending the college relative to the majority student.

Hickman (2009) adopts a similar approach by making the college admission problem into an all-pay auction with incomplete information in order to study the effects of types of affirmative action policies on the racial achievement gap, the enrollment gap, and effort incentives. He finds that, in general, quotas perform better than simple admission preference rules. The reason is that preference rules uniformly subsidize grades without rewarding performance, and therefore have a negative effect on effort incentives. In general, however, the details of the admission rule are important, and the optimal policy depends on parameters, which can only be determined empirically.

In a similar vein, Fryer and Loury (2005) use a tournament model to investigate the categorical redistributions in a winner-take-all market and show that optimally designed tournaments naturally involve "handicapping."[27]

7. EFFICIENCY IMPLICATIONS OF STATISTICAL DISCRIMINATION

In models of statistical discrimination, the use of group identity as a proxy for relevant variables is typically the informationally efficient response of an information-seeking, individually rational agent. Efficiency considerations are therefore especially appropriate in these settings, and a small literature has been devoted to analyzing the different sources of inefficiency arising from statistical discrimination. This is in sharp contrast to Becker-style taste discrimination models where efficiency is not an issue. In models where discrimination arises directly from preferences, any limitation in the use of group identity generates some inefficiencies, at least directly.

7.1 Efficiency in models with exogenous differences

In Phelps' (1972) basic model analyzed in Section 2, discrimination has a purely redistributive nature. If employers were not allowed to use race as a source of information, wages would then equal the expected productivity of the entire population conditional on signal θ. Thus, wage equation (1) is replaced by:

$$E(q|\theta) = \frac{\sigma^2}{\sigma^2 + \sigma_\varepsilon^2}\theta + \frac{\sigma_{\varepsilon^2}}{\sigma^2 + \sigma_{\varepsilon^2}}[\lambda\mu_B + (1-\lambda)\mu_W]$$

where λ is the share of group-B workers in the labor market, $\sigma^2 = \lambda^2\sigma_B^2 + (1-\lambda)^2\sigma_W^2$, and $\sigma_\varepsilon^2 = \lambda^2\sigma_{\varepsilon B}^2 + (1-\lambda)^2\sigma_{\varepsilon W}^2$. Assuming a total population size of 1, total product would be equal to average productivity, $\mu = \lambda\mu_B + (1-\lambda)\,\mu_W$. This quantity is the same as when the employers are allowed to discriminate by race. Thus, there is no efficiency gain from discrimination. This equivalence, however, is an artifact of the extreme simplicity of the model and is not robust to many simple extensions.

Suppose, as an illustration, that there are two jobs in the economy, with different technologies. Assume that workers with productivity less than the population average μ are only productive in job 1, and workers with productivity greater than μ are only productive in job 2. In this case, $E(q\,|\,\mu) = \mu$; therefore, it is optimal for firms to allocate workers with signals $\theta < \mu$ to job 1 and workers with signals $\theta \geq \mu$ to job 2. Some mismatches will occur. If populations have different population averages, $\mu_B \neq \mu_W$, then the optimal allocation rule follows thresholds θ_j, $j \in \{B, W\}$ computed to satisfy

[27] Schotter and Weigelt (1992) found evidence that affirmative action may increase the total output in an asymmetric tournament in a laboratory setting. Calsamiglia, Franke, and Rey-Biel (2009) have similar findings in a real-world field experiment involving school children. See Holzer and Neumark (2000) for a detailed survey of available evidence regarding the incentive effects of affirmative action policies.

$E(q \mid \theta_j) = \mu$, which differ by group. Mismatch increases when employers are not allowed to discriminate by race, because race functions effectively as a proxy for productivity.

When human capital investment is endogenous, as in Lundberg and Startz's (1983) version of Phelps' model, efficiency also depends on the human capital investment cost paid by workers. One source of inefficiency of discriminatory outcomes is that the marginal worker from the dominant group pays a higher cost than the marginal worker from the discriminated group. Using the parameterization presented in Section 2.2.2, the marginal worker produces:

$$MP(X) = a + bX^* = a + \frac{b^2}{c}\frac{\sigma^2}{\sigma^2 + \sigma_{\varepsilon j}^2}$$

(see equation 2) after spending $C(X) = cX^2/2$ in investment costs. Hence the net social product of human capital investment in group-j is:

$$MP(X) - MP(0) - C(X) =$$
$$a + \frac{b^2}{c}\frac{\sigma^2}{\sigma^2 + \sigma_{\varepsilon j}^2} - a - \frac{b^2}{2c}\left(\frac{\sigma^2}{\sigma^2 + \sigma_{\varepsilon j}^2}\right)^2 = \frac{b^2}{c}\left(1 - \frac{1}{2}\frac{\sigma^2}{\sigma^2 + \sigma_{\varepsilon j}^2}\right)$$

To generate a discriminatory equilibrium, assume $\sigma_{\varepsilon B}^2 > \sigma_{\varepsilon W}^2$. In this case it is efficient to transfer some units of training from high cost W workers to low-cost B workers. In general, a ban on the use of race results in a more efficient solution relative to the statistical discrimination outcome.

However, as Lundberg and Startz (1983) note in their conclusion, this result is not robust, and it is meant to illustrate a more general principle that in a second-best world, as one in which there is incomplete information, "there is no reason to assume that approaching the first best—using more information—is welfare improving. Since the problem of incomplete information is endemic in situations of discrimination, considerations of the second best are a general concomitant to policy questions in this area."

Other papers focus therefore on sources for the opposite outcome, that is showing that statistical discrimination may be efficiency enhancing. This depends on the details of the model specification and sometimes on the parameterization of the model.

Schwab (1986), for example, focused on one specific type of mismatching that statistical discrimination generates. In this paper, workers can pool with other workers in a "standardized" labor market in which individual productivity cannot be detected, and therefore everybody is paid a wage equal to the average productivity in the pool of workers. Workers can, alternatively, self-employ and receive compensation that is an increasing function of their ability. The marginal worker is indifferent between self-employment and the standardized market. However, her productivity in the standardized market must be higher than her wage, because all of the workers in her pool have lower productivity. This is an informational externality, which implies an employment level in the standardized market lower than socially optimal.

Consider adding to this model a second group of workers with higher average ability in the standardized market. In an equilibrium with statistical discrimination, wage in the standardized market will depend on group identity, and will be higher for members of the second group. A ban on statistical discrimination practices will equalize such wage, but will have ambiguous effects on efficiency. It will increase standardized market employment for members of the less productive group, therefore approaching the first-best solution for this group, but the opposite happens for members of the more productive group. The total effect depends on the details of the ability distribution in the two groups.[28]

7.2 Efficiency in models with endogenous differences

The same effects play a role in the equilibrium models of statistical discrimination analyzed in Sections 3 and 4: the efficient allocation of workers to jobs, the role of the informational externalities due to imperfect information. In addition, efficiency may depend on the effects on the cost of human capital investment, and, depending on the technology, the role of complementarities in the production function.

Two broad sets of questions can be asked in this context. First, does the planners' problem solution imply differential treatment across groups? Second, are discriminatory equilibria more efficient than symmetric, nondiscriminatory equilibria?

7.2.1 The planners' problem

A comprehensive analysis of the various effects is performed in Norman (2003), where symmetric outcomes are compared to discrimination in the planners' problem.

Norman adopts a simplified version of the model in Moro and Norman (2004) and shows first that if the planner is allowed to discriminate between groups, then the production possibility frontier expands. This is a direct implication of employers' imperfect information. Assume for simplicity there are only two signals, H(igh) and L(ow), such that the probability that a qualified worker receives a high signal is $f > 1/2$, whereas the same probability for a low-signal worker is $(1 - f)$. For an intuition, consider the case where groups have equal size, and compare the situations where both groups invest the same amount π with the case where they invest differently, $\pi_B < \pi_W$, but aggregate investment is equal to π.

It is not difficult to see that the production possibility frontier expands with group inequality. Any factor input combination (C, S) with $S > 0$, $C > 0$ achievable in the symmetric case can be improved upon by replacing a high-signal B worker employed in the complex task with a high-signal W worker employed in the simple task.

[28] A similar model is also analyzed in Haagsma (1993), who considers also the effects of varying labor supply.

Substituting these two workers does not change the input in the simple task, but it increases expected input in the complex task because the expected productivity in the complex task is higher for W workers,

$$\frac{\pi_W f}{\pi_W f + (1-\pi_W)(1-f)} > \frac{\pi_B f}{\pi_B f + (1-\pi_B)(1-f)}. \tag{82}$$

Incomplete information generates misallocation of workers to task. In an asymmetric equilibrium race functions as an additional signal that moderates the informational problem.

However, to generate higher investment in group W the planner has to pay high signal workers from this group a higher premium. Such premium can be "financed" via a transfer or resources from group B, or exploiting the informational efficiency gains. Norman shows with two parametric examples the role of the difference between a linear technology and a technology with complementarities. The crucial result is that when there are complementarities, the discriminatory solution *may* result in Pareto-gains, that is, in an outcome where both groups are better off. On the other hand, when technology is linear, the planner can implement the efficient asymmetric solution only by transferring resources from the discriminated group to the dominant group.

It is possible to illustrate this result with a simple parametric example. Consider a technology given by $y(C,S) = \sqrt{CS}$ with cost of investment equal to 0 for half of the workers of either group, and 0.1 for everybody else. As in the example described above, there are only two feasible signals, H and L, and with $f = 2/3$.

Consider first the situation where the planner is constrained to a symmetric outcome. The advantage of the cost distribution we adopted is that the solution is either $\pi = 1/2$ or $\pi = 1$ so we only need to compare these two cases. When $\pi = 1$ everybody is equally productive in either task, therefore the optimal solution is to assign half the population to each task, and total output is $y = 0.5$. When $\pi = 1/2$, one can easily compute that the optimal solution is to assign all H workers to the complex task and all L workers to the simple task. In this case $C = 2/3 * 1/2$ and $S = 1/2$, which implies $y = 0.5\sqrt{2/3} < 0.5$. Cost of investment is zero when $\pi = 1/2$ and 0.05 when $\pi = 1$. Hence the optimal solution is $\pi = 1$. In this solution, there are 2/3 workers with signal H, hence to implement this outcome, the planner can pay L workers 0 and H workers 3/2. Incentives to invest are 3/2 * (2/3 − 1/3) = 1/2.

To solve for the asymmetric outcome, note that in the symmetric solution 1/2 of the workers are employed in the simple task but do not need to be qualified. Hence, it would be more efficient if we could "tag" half the workers and induce them not to invest in human capital. Using race, the planner can have all W workers replicate what they do in the previous outcome, and all B workers not to invest in human capital. Then, assign all W workers to the complex task and all B workers to the simple task. Output would be the same, but half of the investment costs would be the saved.

This outcome is implementable by paying all B workers 1/2 regardless of their signal, and paying W workers as before. Total wage bill is 1/2 for B workers, and $3/2 * 2/3 * 1/2 = 1/2$ for W workers. Because of the savings in investment cost, the B group is more than fully compensated in this outcome.

What this example shows is that complementarities in the production function coupled with specialization allow the planner to reduce investment cost without changing output. This would be impossible in the linear case because less investment implies lower output. Therefore, the gains from specialization cannot be redistributed across groups without breaking incentive compatibility. In a parametric example, Norman shows that even in the linear case there may be efficiency gains from discrimination in the planners' problem (arising from reduced mismatching), but that the added investment for the dominant group must be supported using transfers from the discriminated group.

7.2.2 The efficiency of discriminatory equilibria

Considering the case of the equilibrium model in Moro and Norman (2004) with a linear technology, where discrimination results from coordination failure (see Section 3). Note that equilibria are Pareto-ranked. To see this, the model with a single group of workers displaying two equilibrium levels of human capital investment, $\pi_1 > \pi_2$. Under π_1, wages as a function of θ are weakly greater than under the lower level of human capital investment π_2. Therefore, all workers that either do not invest or that do invest in both equilibria are better off under the high human capital investment equilibrium because they have higher expected wages, which can be computed using (3) by integrating over the relevant distribution of θ, that is f_q for workers that invest, and f_u for workers that do not invest. There is a set of workers that do not invest under $\pi_{2,}$ but do invest and pay the investment cost under π_1. To see that even these workers are better-off, note that because they choose to invest, it must be that the benefits outweigh the cost, that is, $\int w(\theta,\pi_1)f_q(\theta) - c \geq \int w(\theta,\pi_1)f_u(\theta)$. The left-hand side however must be greater than the expected wage of non-investors under π_2, $\int w(\theta,\pi_2)f_u(\theta)$. Therefore $\int w(\theta,\pi_1)f_q(\theta) - c > \int w(\theta,\pi_2)f_u(\theta)$; that is, even these workers strictly prefer the higher investment equilibrium.

Hence, because of the linearity in production, separability between groups implies that the discriminatory equilibrium is not efficient. When production displays complementarities, because of effects that are similar to the one displayed in the example illustrated in the planners' problem, we conjecture the possibility that group-wide Pareto gains may exist in discriminatory equilibria relative to symmetric equilibria.

8. CONCLUSION

This chapter surveyed the theoretical literature on statistical discrimination and affirmative action stressing the different explanation for group inequality that have been developed from the seminal articles of Phelps (1972) and Arrow(1973), and their policy implications.

In this conclusion, we highlight some areas for potentially fruitful future research. First, as we mentioned in Section 5, we still have a relatively poor theoretical understanding on the evolution of stereotypes, under what conditions do they arise and lead to permanent inequality, and how the stereotypes are affected by supposedly temporary affirmative action policies. There is not yet any study on how affirmative action policies might change the dynamics of the between-group inequalities. Can temporary affirmative action measures indeed lead to between-group equalities, as proclaimed in Supreme Court justices' opinion in 1978 and 1993? Second, most of the existing literature on affirmative action has studied a quite stylized version of the policy, assuming that employers follow quotas set by the policymaker. In practice, however, the policy maker rarely sets clearly defined quotas. In addition, there exist agency issues between the policymaker (the principal) and the decision-makers (the agent). As an example that should be familiar in the academic world, consider the case of a college dean and a research department that place different weights on their concern for academic excellence and faculty racial or gender diversity. How affirmative action policies should be optimally designed in light of such agency issues is also an important question to study.

Finally, this survey has not made much connection between the theoretical models and the small existing empirical literature related to statistical discrimination theories. Most of the empirical literature on racial and gender inequality focuses on measuring inequality after controlling for a number of measurable factors without attempting to attribute the unexplained residuals to a specific source of discrimination.[29] Some articles attempt to test implications of statistical discrimination directly, with mixed evidence. For example, Altonji and Pierret (2001) test dynamic wage implications of statistical discrimination.[30] Another growing literature attempts to use statistical evidence to distinguish statistical discrimination from racial prejudice, particularly regarding racial profiling in highway stops and searches.[31] In surveying the trends of Black-White wage inequality, Neal (2010) finds that returns to schooling and other test scores are higher for minorities, evidence that he claims to be counterfactual to statistical discrimination theories based on endogenous differential incentives to acquire skills.[32] However, the

[29] Most of these articles assume or suggest that the unexplained differences should be attributed to racial bias. Interested readers should consult the surveys by Altonji and Blank (1999) and Holzer and Neumark (2000).

[30] See also Lange (2007).

[31] See, e.g., Knowles, Persico and Todd (2001), and Anwar and Fang (2006) for evidence on police racial profiling. Fang and Persico (2010) provide a unified framework to distinguish racial prejudice from statistical discrimination that is applicable in many settings.

[32] For additional evidence on returns to aptitude test scores, see Neal and Johnson (1996) and, with more recent data, Fadlon (2010). See also Heckman, Lochner and Todd (2006) for evidence on returns to education controlling for selection bias.

human-capital-based theories that originate from Arrow's (1973) insight depends crucially on *unobserved* human capital investment; therefore, they do not directly imply that returns to *observable* human capital, such as education, should be different or higher for the dominant group. For example, conditional on education, statistical discrimination can predict that members of the discriminated group exert lower learning effort because they have fewer incentives to do so; but returns to schooling might be higher for them. In addition, the theory only predicts that groups have different returns to the skill signals that are *observed by employers*, not to signals observed by the investigator. Even if we interpret education (or any other observable test score) as a signal of skill, a regression of wages on such signals produces estimates that suffer from omitted variable bias whenever firms also use privately observed signals. The size of this bias depends on group fundamentals in ways that might confuse the inference made by the econometrician.[33]

Nevertheless, we believe that studying ways to reconcile empirical facts about wage differences and the typical theoretical predictions of statistical discrimination theories could be a fruitful area of future research. Some attempts at structurally estimating statistical discrimination models find that even stylized versions of these models fit the data quite well. For example, Moro (2003) structurally estimates a model based on Moro and Norman (2004) using Current Population Survey data and finds that adverse equilibrium selection did not play a role in exacerbating wage inequality during the last part of the 20th century. Fang (2006) estimates, using Census data, an equilibrium labor market model with endogenous education choices based on Fang (2001) to assess the relative importance of human capital enhancement versus ability signaling in explaining the college wage premium. Bowlus and Eckstein (2002) estimate a structural equilibrium search model to distinguish the roles of skill differences among groups and employers' racial prejudice to explain racial wage inequality.[34] However, these estimates are not designed to perform model validation. Research addressing the identification issue of how to disentangle different sources of group inequality (being from statistical, taste-based discrimination, or from differences in groups' fundamentals) would be especially welcome.

REFERENCES

Abdulkadiroglu, A., 2005. College Admissions with Affirmative Action. International Journal of Game Theory 33, 535–549.

Aigner, D., Cain, G., 1977. Statistical Theories of Discrimination in the Labor Market. Ind. Labor Relat. Rev. 30 (2), 175–187.

Altonji, J., Blank, R., 1999. Race and Gender in the Labor Market. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics, vol. 3C Elsevier, North Holland, pp. 3143–3259 (Chapter 48).

[33] Moro and Norman (2003b) show this point formally.

[34] See also Flabbi (2009) for the case of gender wage differences in a model with both matching and wage bargaining.

Altonji, J., Pierret, C.R., 2001. Employer Learning and Statistical Discrimination. Q. J. Econ. 116 (1), 313–350.
Antonovics, K., 2006. Statistical Discrimination and Intergenerational Income Mobility. mimeo, UC San Diego.
Anwar, S., Fang, H., 2006. An Alternative Test of Racial Profiling in Motor Vehicle Searches: Theory and Evidence. Am. Econ. Rev. 96 (1), 127–151.
Arcidiacono, P., Aucejo, E., Fang, H., Spenner, K.I., 2009. Does Affirmative Action Lead to Mismatch? A New Test and Evidence. NBER Working Paper No. 14885.
Arcidiacono, P., 2005. Affirmative Action in Higher Education: How do Admission and Financial Aid Rules Affect Future Earnings? Econometrica 73 (5), 1477–1524.
Arrow, K.J., 1973. The Theory of Discrimination. In: Ashenfelter, O., Rees, A. (Eds.), Discrimination in Labor Markets. Princeton University Press, pp. 3–33.
Ayres, I., Brooks, R., 2005. Does Affirmative Action Reduce the Number of Black Lawyers? Stanford Law Rev. 57 (6), 1807–1854.
Barnes, K.Y., 2007. Is Affirmative Action Responsible for the Achievement Gap Between Black and White Law Students? Northwest. Univ. Law Rev. 101 (4), Fall, 1759–1808.
Becker, G.S., 1957. The Economics of Discrimination. University of Chicago Press, Chicago.
Blume, L., 2006. The Dynamics of Statistical Discrimination. Econ. J. 116, F480–F498.
Cain, A., 1986. The Economic Analysis of Labor Market Discrimination: A Survey. In: Ashenfelter, O., Layard, R. (Eds.), Handbook of Labor Economics, vol. 1 Amsterdam, North Holland, pp. 693–785 (Chapter 13).
Calsamiglia, C., Franke, J., Rey-Biel, P., 2009. The Incentive Effects of Affirmative Action in a Real-Effort Tournament. Unpublished Working Paper, Universitat Autonoma de Barcelona.
Chambers, D.L., Clydesdale, T.T., Kidder, W.C., Lempert, R.O., 2005. The Real Impact of Eliminating Affirmative Action in American Law Schools: An Empirical Critique of Richard Sander's Study. Stanford Law Rev. 57 (6), 1855–1898.
Chan, J., Eyster, E., 2003. Does Banning Affirmative Action Lower College Student Quality? Am. Econ. Rev. 93 (3), 858–872.
Chaudhuri, S., Sethi, R., 2008. Statistical Discrimination with Peer Effects: Can Integration Eliminate Negative Stereotypes? Rev. Econ. Stud. 75, 579–596.
Chung, K.S., 1999. Affirmative Action as an Implementation Problem. Unpublished Working Paper, University of Minnesota.
Chung, K.S., 2000. Role Models and Arguments for Affirmative Action. Am. Econ. Rev. 90 (3), 640–648.
Coate, S., Loury, G., 1993a. Will Affirmative Action Eliminate Negative Stereotypes? Am. Econ. Rev. 83 (5), 1220–1240.
Coate, S., Loury, G., 1993b. Antidiscrimination Enforcement and the Problem of Patronization. American Economic Review Papers and Proceedings 83 (2), 92–98.
Cornell, B., Welch, I., 1996. Culture, Information, and Screening Discrimination. J. Polit. Econ. 104 (3), 542–571.
DeGroot, M.H., 2004. Optimal Statistical Decisions, first ed. Wiley-Interscience.
Eeckhout, J., 2006. Minorities and Endogenous Segregation. Rev. Econ. Stud. 73 (1), 31–53.
Bowlus, A., Eckstein, Z., 2002. Discrimination and Skill Differences in an Equilibrium Search Model. Int. Econ. Rev. 43 (4), 1309–1345.
Epple, D., Romano, R., Sieg, H., 2002. On the Demographic Composition of Colleges and Universities in Market Equilibrium. American Economic Review Papers and Proceedings 92 (2), 310–314.
Epple, D., Romano, R., Sieg, H., 2008. Diversity and Affirmative Action in Higher Education. J. Public Econ. Theory 10 (4), 475–501.
Fadlon, Y., 2010. Statistical Discrimination and the Implications of Employer-Employee Racial Matches. Unpublished manuscript, Vanderbilt University.
Fang, H., 2001. Social Culture and Economic Performance. Am. Econ. Rev. 91 (4), 924–937.
Fang, H., Loury, G., 2005a. Toward An Economic Theory of Dysfunctional Identity. In: Christopher, B (Ed.), Social Economics of Poverty: On Identities, Groups, Communities and Networks. Routledge, Barrett, London, pp. 12–55.

Fang, H., Loury, G., 2005b. Dysfunctional Identities Can Be Rational. American Economic Review Papers and Proceedings 95 (2), 104–111.
Fang, H., Norman, P., 2006. Government-mandated discriminatory policies. Int. Econ. Rev. 47 (2), 361–389.
Fang, H., 2006. Disentangling the college wage premium: estimating a model with endogenous education choices. Int. Econ. Rev. 47 (4), 1151–1185.
Fang, H., Persico, N., 2010. Distinguishing prejudice from statistical discrimination: a unified framework. mimeo, University of Pennsylvania.
Flabbi, L., 2009. Gender Discrimination Estimation in a Search Model with Matching and Bargaining. Int. Econ. Rev. forthcoming.
Fryer, R., 2007. Belief flipping in a dynamic model of statistical discrimination. J. Public Econ. 91 (5–6), 1151–1166.
Fryer, R., Loury, G., 2005. Affirmative Action in Winner-Take-All Markets. Journal of Economic Inequality 3 (3), 263–280.
Fryer, R., Loury, G., 2008. Valuing Identity: The Simple Economics of Affirmative Action Policies. unpublished manuscript.
Fu, Q., 2006. A Theory of Affirmative Action in College Admissions. Econ. Inq. 44, 420–428.
Haagsma, R., 1993. Is Statistical Discrimination Socially Efficient? Inf. Econ. Policy 5, 31–50.
Heckman, J., Lochner, L., Todd, P., 2006. Earnings Functions, Rates of Returns and Treatment Effects: The Mincer Equation and Beyond. In: Hanusheck, E., Welch, F. (Eds.), Handbook of the Economics of Education, vol. I. North-Holland, pp. 307–458 (Chapter 7).
Hickman, B.R., 2009. Effort, Race Gaps and Affirmative Action: A Game-Theoretic Analysis of College Admissions. Unpublished, University of Iowa.
Ho, D.E., 2005. Why Affirmative Action Does Not Cause Black Students to Fail the Bar. Yale Law J. 114 (8), 1997–2004.
Holden, S., Rosen, A., 2009. Discrimination and Employment Protection. CESifo Working Paper Series No. 2822.
Holzer, H., Neumark, D., 2000. Assessing Affirmative Action. J. Econ. Lit. 38 (3), 483–568.
Knowles, J., Persico, N., Todd, P., 2001. Racial Bias in Motor Vehicle Searches: Theory and Evidence. J. Polit. Econ. 109, 203–228.
Lange, F., 2007. The speed of employer learning. J. Labor Econ. 25 (1), 651–691.
Levin, J., 2009. The Dynamics of Collective Reputation,. BE Journal of Theoretical Economics 9 (1), August 2009.
Lang, K., 1986. A Language Theory of Discrimination. Q. J. Econ. 101 (2), 363–382.
Loury, L.D., Garman, D., 1995. College Selectivity and Earnings. J. Labor Econ. 13 (2), 289–308.
Lundberg, S., 1991. The Enforcement of Equal Opportunity Laws under Imperfect Information: Affirmative Action and Alternatives. Q. J. Econ. 106, 309–326.
Lundberg, S., Startz, R., 1983. Private Discrimination and Social Intervention in Competitive Markets. Am. Econ. Rev. 73 (3), 340–347.
Mailath, G., Samuelson, L., Shaked, A., 2000. Endogenous Inequality in Integrated Labor Markets with Two-Sided Search. Am. Econ. Rev. 90 (1), 46–72.
Moro, A., 2003. The effect of statistical discrimination on black-white wage in-equality: estimating a model with multiple equilibria. Int. Econ. Rev. 44 (2), 457–500.
Moro, A., Norman, P., 1996. Affirmative Action in a Competitive Economy. CARESS Working Paper #96–08, University of Pennsylvania.
Moro, A., Norman, P., 2003a. Affirmative Action in a Competitive Economy. J. Public Econ. 87 (3–4), 567–594.
Moro, A., Norman, P., 2003b. Empirical Implications of Statistical Discrimination on the Returns to Measures of Skills. Ann. Econ. Stat. 71–72, 399–417.
Moro, A., Norman, P., 2004. A General Equilibrium Model of Statistical Discrimination. J. Econ. Theory 114 (1), 1–30.
Neal, D., 2010. Black–white labour market inequality in the United States. In: Durlauf, S.N., Blume, L.E. (Eds.), The New Palgrave Dictionary of Economics. Palgrave Macmillan 2008, The New Palgrave

Dictionary of Economics Online, Palgrave Macmillan. 23 February 2010, DOI: 10.1057/9780230226203.0139.

Neal, D., Johnson, W., 1996. The role of pre-market factors in black-white wage differences. J. Polit. Econ. 104, 869–895.

Norman, P., 2003. Statistical Discrimination and Efficiency. Rev. Econ. Stud. 70 (3), 615–627.

Phelps, E., 1972. The statistical theory of racism and sexism. Am. Econ. Rev. 62, 659–661.

Rosén, Å., 1997. An Equilibrium Search-Matching Model of Discrimination. Eur. Econ. Rev. 41 (8), 1589–1613.

Rothstein, J., Yoon, A., 2008. Mismatch in Law School. Unpublished manuscript, Princeton University.

Sander, R., 2004. A Systemic Analysis of Affirmative Action in American Law Schools. Stanford Law Rev. 57 (2), 367–483.

Sattinger, M., 1998. Statistical Discrimination with Employment Criteria. Int. Econ. Rev. 39 (1), 205–237.

Schotter, A., Weigelt, K., 1992. Asymmetric Tournaments, Equal Opportunity Laws, and Affirmative Action: Some Experimental Results. Q. J. Econ. 107 (2), 511–539.

Schwab, S., 1986. Is Statistical Discrimination Efficient? Am. Econ. Rev. 76 (1), 228–234.

Spence, M., 1974. Job Market Signaling. Q. J. Econ. 87, 355–374.

Stiglitz, J., 1973. Approaches to the Economics of Discrimination. American Economic Review Papers and Proceedings 63 (2), 287–295.

Verma, R., 1995. Asymmetric information, search theory and some commonly observed phenomena. University of Pennsylvania Ph.D. dissertation.