Affirmative Action in the Presence of Income Heterogeneity

Guha, Brishti and Roy Chowdhury, Prabal

Jawaharlal Nehru University, Indian Statistical Institute, New Delhi

2017

Affirmative Action in the Presence of Income Heterogeneity

Brishti Guha* Prabal Roy Chowdhury†

Abstract

We examine affirmative action, class-based (CAA), as well as identity-based (IAA), in an economy with income heterogeneity and diverse identity groups. We establish that with class-based affirmative action there exists an equilibrium that is colour-blind, i.e. the poor face the same standards while being assigned to skilled jobs irrespective of their identity, as do the rich. Further, if affirmative action is identity-based, then there exists an equilibrium that has elements of patronization in that black workers of both income classes face lower standards relative to their white counterparts. Interestingly, a comparison shows that the relative preferences for the two policies are completely class-based, with all poor workers preferring CAA, whereas all rich workers prefer IAA. In fact, poor black workers prefer CAA though they benefit from affirmative action under both IAA and CAA, whereas the rich white workers prefer IAA though they are not protected under either form of affirmative action. Moreover, there exist economies where CAA ensures that all poor workers are assigned to the skilled task with positive probability, whereas under IAA that is not necessarily true.

JEL No.: J71, J78, D78, D82.

Key Words: Income heterogeneity; identity; affirmative action; class-based affirmative action; identity-based affirmative action.

*School of International Studies, Jawaharlal Nehru University, New Delhi, India; E-mail: brishtiguha@gmail.com

†Economics and Planning Unit, Indian Statistical Institute, New Delhi, India; E-mail: prabalrc@isid.ac.in

1 Introduction

Both identity and poverty can, and have historically contributed to, certain sections of society being disadvantaged. Income heterogeneity is therefore of salience when examining affirmative action for the disadvantaged. Thus in this paper we examine affirmative action,[1] both class, as well as identity-based, in a society that is divided along lines of class, as well as identity.

While affirmative action has traditionally been identity-based, class-based affirmative action is being debated, as well as implemented in several countries. Reece (2011) describes how, in the US, affirmative action in government contracting has been moving away from race-based preferential treatment to that based on economic criteria. Programs like HUBZone and TACPA award contracts preferentially to small business owners in socio-economically disadvantaged areas, whose employees are mainly economically disadvantaged youth (Reece 2011). Moreover, many countries that have affirmative action in college admissions give at least some weight to income. Examples include Brazil,[2] France,[3] Israel,[4] New Zealand[5] and South Africa.[6] The reservation policy in India for example, while predominantly based on caste, does take some account of class in that there is a cutoff income so that any potential 'Other Backward Caste (OBC)' beneficiary with income exceeding the cutoff is excluded.[7] In fact in the Indian context Thomas Piketty argues that "... the long-term objective should be to gradually move away from a caste-based reservation system to

[1]Many countries adopt affirmative action policies in favour of traditionally under-privileged groups, with the list of countries practicing some form of affirmative action including, among others, the US, Canada, India, South Africa, New Zealand, Malaysia, China, Israel, and Sri Lanka. Affirmative action has been used since the 1960s in the US, with race and sex being part of hiring criteria. Similarly the Canadian Employment Equity Act gives preferences to women, aboriginals, and minorities. In India, caste is a basis for reservation in government jobs and publicly funded educational institutions. In South Africa, the Employment Equity Act of 1998 mandated that companies with more than 50 employees must meet proportional quotas in their hiring of disadvantaged groups. In New Zealand affirmative action has given those of Maori descent better access to university and financial aid since 1993. Brazil has quotas for racial minorities and the poor in higher education.

[2]In Brazil, quotas for the poor co-exist with those for racial minorities in federal universities and in some civil service jobs (the Economist, April 26, 2013). It is attempting to implement an affirmative action policy in college admissions that is both race-conscious and class-based. In particular the admission policy is considering whether applicants have public school backgrounds and are from poor families, while also ensuring a certain racial mix (see http://www.universityworldnews.com/article.php?story=20160419134316317).

[3]Highly-ranked French schools are required to maintain quotas for students from poorer families (Le Monde, December 17, 2008).

[4]In Israel four elite universities implemented a programme of class-based affirmative action that ignored racial and ethnic criteria (Alon 2011). At the same time, from 2008 onwards, a proportion of seats in the Israeli civil service was reserved for Arabs (Haaretz, 2nd April, 2010).

[5]In New Zealand, some universities which traditionally had quotas for students of Maori ethnicity began considering giving preference to economically disadvantaged students as well, regardless of ethnicity (see http://mauistreet.blogspot.in/2013/06/affirmative-action-class-or-ethnicity.html).

[6]In South Africa, universities in Cape Town have started to consider whether children of impoverished black families should be treated the same way as children of more affluent blacks, and have begun to incorporate both racial and family background considerations in their admission policies (see http://www.universityworldnews.com/article.php?story=20160419134316317).

[7]In India, the National Commission for Backward Classes has proposed a ceiling of Rs 10.5 lakh in 2015.

a system of reservation that is more based on parental income, parental wealth."[8]

We examine an economy which is divided along lines of both race, "white" and "black", and class, "rich" and"poor". Thus, potential workers may belong to any of four possible categories - they may be poor blacks, poor whites, rich blacks, or rich whites. We also assume that on the average blacks are poorer relative to the whites. This is realistic because socially marginalized communities, who are usually the intended beneficiaries of identity-based affirmative action, have also traditionally been at an economic disadvantage. Following Arrow (1972) and Coate and Loury (1993), we examine a model where a representative firm can observe both a worker's identity as well as her class, but not her skill level, regarding which they receive an imperfect signal. Further, the poor have higher opportunity costs of investing in education relative to the rich. This may arise because of various reasons that we discuss later in the paper. The belief of the firm regarding any worker is conditioned both by their signal, as well as their perception regarding the average skill level of the concerned group. Based on this belief firms assign workers to either a skilled task (with higher wages), or to an unskilled one. The workers, in turn, may or may not acquire the requisite skill depending on their own costs of doing so, as well as on the belief that the employers have about their group.

We start by analysing equilibria in the absence of any affirmative action policies, either class, or identity based, finding that there can be multiple equilibria. We start by identifying three equilibria that are salient to our analysis. Moreover, all three equilibria share certain properties. First, stereo-typing exists, at least for certain income classes. Moreover, such stereo-typing takes the most extreme form (in a sense formalized later) whenever it exists. Second these equilibria are non-trivial in that at least one of the groups acquires a positive level of skill. Where the equilibria differ are the implicit assumptions regarding the way income levels may or may not affect stereo-typing. The first equilibrium formalizes the idea that poverty can be such a burden that the poor acquire no education at all, so that the question of stereo-typing does not arise for them. For the rich (think of the not-so-poor), stereo-typing is however very much present. The second equilibrium assumes that stereo-typing is present at all levels of income. The third equilibrium assumes that while the poor face stereo-typing, the rich blacks escape it.

Before analysing the two forms of affirmative action, we establish some preliminary results that will be used while comparing worker and firm utility across various equilibria. In particular we find that workers' utility is inversely related to the cutoffs that workers of their category face for assignment to the skilled task. Further, firm profit is positively related to skill acquisition by all four groups.

We find that if class-based affirmative action (CAA) is in place, then there exists an equilibrium that is colour-blind in the sense that assignment to skilled jobs depends on class, but not on identity, and moreover, while being assigned to the skilled task the poor face less stringent standards relative to the rich. How does the imposition of CAA affect the workers? Some broad patterns emerge. We find that irrespective of which one of the three equilibria described earlier prevails, all poor workers are better off under CAA. Turning to the effect on black workers, we find that rich black workers would be worse off under CAA in case they

[8]See http://www.livemint.com/Politics/sAAIiM843JTZDuTCzZWpZO/Thomas-Piketty–India-has-a-very-strong-legacy-of-extreme-i.html.

have already escaped stereo-typing. Otherwise, black workers benefit irrespective of their class.

We then characterize equilibrium under identity-based affirmative action (IAA), finding that there exists an equilibrium where black workers are subject to less stringent standards while being assigned to the skilled task, thus exhibiting elements of patronization as developed in Coate and Loury (1993). Intuitively, given that acquiring education is costlier for the poor, and that blacks are poorer on the average, the firm believes, correctly, that black workers are less likely to get skilled. Thus equalizing the number of black and white workers in the skilled task is difficult, forcing firms to apply less stringent standards to black workers. Turning to the effect of IAA, we find that black workers gain under IAA irrespective of the prevailing equilibrium, while rich white workers have a lower utility. Even poor white workers are worse off unless poverty was a severe constraint in the original equilibrium.

We next turn to comparing the color-blind equilibrium under CAA with the patronising equilibrium under IAA. We show that the workers' preference for either IAA, or CAA is divided along class-lines, with both rich black and rich white workers preferring identity-based affirmative action, while all poor workers prefer a class-based policy. Rich black workers, who have favored status under IAA but not under CAA, prefer identity-based affirmative action, whereas poor white workers, who enjoy favored status only under CAA, prefer CAA to IAA.

Interestingly, poor black workers prefer CAA though they benefit from both forms of affirmative action, while rich white workers prefer IAA though they don't benefit from either. This result can again be traced to the fact that the poor, who find acquiring education costlier, are less likely to do so. Hence, given that black workers are heterogenous in terms of income and also include rich members who are more likely to get skilled, equalization of representation under IAA, where the concerned groups are the black and white workers, is relatively easier relative to that under CAA where the concerned groups are the poor and the rich. The shadow price of equality is therefore higher under CAA. Consequently, the implicit subsidy obtained by the firm from assigning a poor black worker to the skilled task is higher under a class-based policy, and thus it offers a less stringent test under a class-based policy. Similarly, the implicit tax on the firm from assigning a rich white worker to the skilled task is higher under a class-based policy. Hence the result.

We finally consider an economy such that in the absence of any affirmative action the poor, irrespective of colour, acquire no skill and are never assigned to task 1, so that one issue of interest is to find a policy that benefits all poor workers, irrespective of identity. We argue that if the economy is such that inequality among blacks is severe, and moreover neither the blacks nor the whites are under-represented among the poor, then under any CAA equilibrium (barring the trivial one) all poor workers are better off, while IAA may leave poor white workers untouched.

1.1 Literature Review

The literature closest to ours is the one on statistical discrimination, pioneered by Phelps (1972) and Arrow (1973).[9] The central idea behind statistical discrimination is that individ-

[9]The literature on bias/discrimination has also explored some related ideas. Becker (1957) develops a theory of employment discrimination grounded in preference-based discrimination. In related work, Welch

ual attributes are not perfectly observable, so that employers use group attributes in making their decisions. The idea of statistical discrimination was further developed by Coate and Loury (1993), who use it to analyze the effect of affirmative action on stereo-types. Most strikingly, they find that affirmative action can lead to patronization whereby the target group can be held to a lower standard vis-á-vis the non-target group. Fang and Moro (2010) provides an excellent survey which incorporates more recent theories of discrimination, including not just those driven by coordination failures (as in Coate and Loury, 1993), but also by other factors such as informational externalities (e.g. Moro and Norman, 2004).

Turning to the theoretical literature,[10] Moro and Norman (2003) examine affirmative action in a framework where wages are endogenously determined, finding that affirmative action may improve the investment level by the discriminated workers in the worst equilibrium. Fryer (2007) examines firms with a hierarchical labour structure, showing that if an employer discriminates against a group of workers in her initial hiring, she may actually favor successful members of that group when she promotes from within the firm. Fang and Norman (2006) show that in the presence of racial discrimination in public sector jobs, members of discriminated groups may be better off in that they acquire a greater level of investment. Lundberg (1991) and Lundberg and Strotz (1983) both examine the impact of affirmative action when regulators do not observe the firms' personnel policies, so that there is worse information about minorities. Both find that affirmative action improves efficiency in human capital investment, but reduces efficiency in production, so that the net effect is ambiguous. Fryer and Loury (2013) examines the effect of group identity being observable on the efficacy of diversity-enhancing policies.

The present paper contributes to this literature by analyzing and comparing two forms of affirmative action in a society divided along both race and class lines, where the poor face higher costs of investing in skill acquisition. We then use this framework to compare and contrast identity based affirmative action policies with income based ones.

Another strand of the literature examines affirmative action in school choice using a mechanism design approach. Abdulkadiroğlu and Sönmez (2003) introduce the student-proposing deferred acceptance algorithm and the top-cycles algorithm, arguing that these algorithms perform better relative to the ones being adopted in practice. Relatedly, Abdulkadiroğlu (2005) examines college admissions when colleges have preferences rather than given priorities. Some papers examine specific assignment protocols, with Westkamp (2013) examining the German University admissions systems, whereas Kamada and Kojima (2011) study the Japanese Residency Matching Program. Further, Kojima (2012) and Hafalir et al. (2013) examine some possible issues with affirmative action, arguing that majority ceilings in some schools can increase competition in other schools, thereby decreasing minority welfare overall.[11] The present paper is clearly different from this strand in that we do not

(1976) studies sector specific employment quotas in a taste-based framework. Theories of bias can also be based on perception, e.g. Bertrand and Mullainathan (2004), and Banerjee et al. (2009).

[10]There is a large empirical literature on affirmative action which has examined, among other issues, the effect of affirmative action on improving black-white earning disparity. One can mention, among many others, Leonard (1984), Smith and Welch (1984), Welch (1989), etc. In the Indian context, one can mention Hnatkovska et al. (2012) and Deshpande and Ramachandran (2014), among others. We refer the readers to Holzer and Neumark (2000) for an analytical survey of both the empirical and the theoretical lietarture.

[11]We refer the readers to Pathak (2011) for a survey.

adopt a mechanism design approach, instead solving for the perfect Bayesian equilibrium of a dynamic game with both asymmetric information and moral hazard.

A third strand of the literature examines the effect of externality. In particular, Loury (1977, 1981) argues that racial segregation along with initial unequal wealth can lead to perpetual differences in economic outcomes for the minority group. Further, in such a situation equal access to opportunities may, or may not be enough in removing inequality depending on the nature of segregation. In our paper however, we completely abstract from such dynamic issues.

The rest of the paper is organized as follows. In the next section we describe the framework of our model. Section 3 examines the outcome under class-based affirmative action. While section 4 considers identity-based affirmative action, in section 5 we examine the implications of replacing identity-based with class-based affirmative action. Finally, Section 6 concludes. The appendix contains some proofs.

2 The Framework

We consider an economy consisting of one representative firm,[12] as well as a unit mass of workers, with the workers being divided along both identity and class lines. Workers may be either black or white (B or W), and either rich or poor (R or P). We thus have four categories of workers, with the number in each category being λ_{ij} where i denotes identity and can take values B or W, while j denotes class and can take values R or P, with $\sum_{i,j} \lambda_{ij} = 1$. Let $\mu_{BR} = \frac{\lambda_{BR}}{\lambda_{BR}+\lambda_{BP}}$ denote the proportion of the rich among the blacks in the population, and $\mu_{WR} = \frac{\lambda_{WR}}{\lambda_{WR}+\lambda_{WP}}$ denote the proportion of the rich among the white workers. We can similarly define μ_{BP} and μ_{WP}, so that $\mu_{BP} + \mu_{BR} = 1$ and $\mu_{WP} + \mu_{WR} = 1$. We shall use the notation $\lambda_i, i = B, W$ to denote $\lambda_{iR} + \lambda_{iP}$, and the notation $\lambda_j, j = P, R$ to denote $\lambda_{Bj} + \lambda_{Wj}$. All agents, workers, as well as the firm, are risk neutral.

The firm assigns workers to one of two tasks, 1 or 2, where task 1 requires skill, whereas task 2 does not. In task 1, the payoff of the firm is x_q (> 0) if the worker is skilled, and $-x_u$ (< 0) otherwise. Further, task 1 carries a positive wage of w, where $w < x_q$. In task 2 on the other hand, both the wages and returns are normalised to zero. One can also interpret an assignment to task 1 as the firm employing the worker, and an assignment to task 2 as the firm not employing the worker at all.

For any worker, acquiring the requisite skill is costly, where this cost depends on both her income level, as well as her type c, where $c \in [0, \infty)$ is idiosyncratic, with a higher c denoting a less talented worker. Consider group {i,j}. The number of ij-type workers whose type is c or less is given by $\lambda_{ij} G(c)$, where $G(c)$ is a continuously differentiable distribution function with density $g(c)$. For a rich worker of type c, her cost of skill acquisition is taken to be c as well. The cost of skill acquisition for a poor worker of type c is however higher at cm, where $m > 1$. Such higher costs could possibly reflect the fact that a poor worker getting skilled may be forced to either give up working altogether, or take up a relatively low paying job during this skill acquisition period, as either taking up a job at all, or taking

[12] Alternatively we can say there is a unit mass of identical firms, with some minor modifications in the formulation.

up jobs with higher pay may not leave enough time to get skilled. In either case there will be a loss of income, something that may be relatively costly for the poor (as it makes consumption smoothing harder).[13] Alternatively, this could arise because with imperfect credit markets, the poor may face higher interest rates in the credit market if they want to fund skill acquisition. Finally, it could be because her environment adversely affects a poor worker's ability to get skilled.

The firm can observe both the identity and the class of the worker assigned to it, but cannot observe whether the worker has acquired the necessary skill. The firms however do observe a signal regarding skill. The signal s, where $s \in [0, 1]$, has distribution $F_q(s)$ if the agent is *q*ualified (i.e. acquired the requisite skill), and $F_u(s)$ if the worker is *u*nqualified. Both $F_q(s)$ and $F_u(s)$ are twice continuously differentiable so that the associated density functions, $f_q(s)$ and $f_u(s)$ respectively, are well defined and continuous for all s. Further, let

$$\phi(s) = \frac{f_u(s)}{f_q(s)},$$

be well defined for all s, and positive for all $0 < s < 1$. The signal is informative in that a higher s signals that the agent is more likely to be qualified. This is formalised as

Assumption 1. $\phi(s)$ *satisfies the monotone likelihood ratio property (henceforth MLRP), i.e.* $\phi(s)$ *is decreasing in* s*. Further, it satisfies the Inada conditions* $\lim_{s\to 0} \phi(s) = \infty$ *and* $\lim_{s\to 1} \phi(s) = 0$.

Let $\rho_{ij}(s_{ij})$ denote the probability that a randomly drawn worker with identity i and class j is assigned to task 1, given that the firm employs a cutoff strategy s_{ij}, assigning workers of this category to task 1 only if they emit signals at least as large as the cutoff. Formally we therefore have that

$$\rho_{ij}(s_{ij}, \pi_{ij}) = \pi_{ij}[1 - F_q(s_{ij})] + (1 - \pi_{ij})[1 - F_u(s_{ij})], \tag{1}$$

where π_{ij} denotes the fraction of workers from the relevant category who gets educated.

We then specify the utility function of the firm, as well as the workers. The payoff to the firm from a worker assigned to task 1 equals

$$\begin{cases} x_q - w, \text{ if the worker is skilled,} \\ -x_u - w, \text{ otherwise,} \end{cases}$$

and equals zero if the worker is assigned to task 2.

Next consider the utility of a worker. Letting c_j denote the cost of acquiring education

[13]Carter and Lybbert (2012), for instance, show how financially constrained households in Burkina Faso could not smooth consumption during periods of income loss.

for a class j worker of type c, so that $c_R = c$ and $c_P = cm$, the utility of such a worker is

$$\begin{cases} w - c_j, \text{ if she is skilled and assigned to task 1,} \\ -c_j, \text{ if she is skilled and assigned to task 2,} \\ w, \text{ if she is unskilled and assigned to task 1,} \\ 0, \text{ otherwise.} \end{cases}$$

The timeline is as follows. In stage 1, the firm and the workers move simultaneously, with the workers deciding on whether to invest or not, and the firm deciding on its task assignment rule. In stage 2, for every worker, the firm observes her identity and income level, as well as the signal from that worker, but not whether she is skilled or not. Finally, in stage 3, the firm decides on task allocation.

Next, we discuss the affirmative action constraints under (i) class-based affirmative action (CAA) and (ii) identity-based affirmative action (IAA).

A *class-based affirmative action* policy (CAA) mandates that the proportion of poor workers assigned to the skilled task must be the same as the corresponding proportion for rich workers, so that

$$\frac{\lambda_{BP}}{\lambda_P}\rho_{BP} + \frac{\lambda_{WP}}{\lambda_P}\rho_{WP} = \frac{\lambda_{BR}}{\lambda_R}\rho_{BR} + \frac{\lambda_{WR}}{\lambda_R}\rho_{WR}. \tag{2}$$

An *identity-based affirmative action* policy (IAA) mandates that the proportion of workers assigned to the skilled task, i.e. task 1, be the same for the black and the white workers, i.e.

$$\begin{aligned} & \frac{\lambda_{BP}}{\lambda_B}\rho_{BP} + \frac{\lambda_{BR}}{\lambda_B}\rho_{BR} = \frac{\lambda_{WP}}{\lambda_W}\rho_{WP} + \frac{\lambda_{WR}}{\lambda_W}\rho_{WR} \\ \Leftrightarrow \quad & \mu_{BP}\rho_{BP} + \mu_{BR}\rho_{BR} = \mu_{WP}\rho_{WP} + \mu_{WR}\rho_{WR}. \end{aligned} \tag{3}$$

Finally, let $P(s'_{ij}, \pi_{ij})$ denote the firm's expected payoff from assigning a worker belonging to group {i,j} who emits a signal exceeding s'_{ij} to task 1 when it believes that a fraction π_{ij} of workers from this group got educated:

$$P(s'_{ij}, \pi_{ij}) = \pi_{ij}[1 - F_q(s'_{ij})](x_q - w) - (1 - \pi_{ij})[1 - F_u(s'_{ij})](x_u + w). \tag{4}$$

3 Equilibrium in the Absence of Affirmative Action

We first analyze the benchmark case when there is no affirmative action. Note that in this case the outcomes for the four groups can be examined separately.

The firms' decisions: Suppose the firm faces a worker of group $\{i, j\}$, $i \in \{W, B\}$, $j \in \{P, R\}$, who emits a signal s. If the firm believes that a proportion π of the workers in this group are skilled, then the firm's belief that this particular worker is skilled, conditional on the signal

s, is given by

$$B(\pi, s) \equiv \frac{\pi f_q(s)}{\pi f_q(s) + (1-\pi) f_u(s)}. \tag{5}$$

Hence the firm assigns this worker to task 1 if and only if the expected profits from doing so exceed the profits from assigning her to task 2 (which is normalised to zero), i.e. $B(\pi, s)(x_q - w) - (1 - B(\pi, s))(x_u + w) \geq 0$, i.e.

$$z \equiv \frac{x_q - w}{x_u + w} \geq \frac{1-\pi}{\pi} \phi(s). \tag{6}$$

Given assumption 1, the firm's decision is characterised by a cutoff s_{ij} such that all workers with a signal greater than s_{ij} are assigned to task 1, where s_{ij} solves (6) with equality. Given that from MLRP it follows that s_{ij} is monotonic (decreasing) in π, $s_{ij}(\pi)$ is well defined. Moreover, note that $s_{ij}(\pi)$ does not depend on the identity of the group $\{i, j\}$, and can thus be denoted as $s(\pi)$. Consequently the graph of $s(\pi)$, call it EE, where EE is a mnemonic for *e*mployer (i.e. the firm), is negatively sloped in $s - \pi$ space (see Figure 1).

The workers' decision: Next consider the decision problem facing a group $\{i, j\}$ worker of type c who believes that the firm will assign her to task 1 if and only if she emits a signal that is at least as high as s_{ij}. Ignoring the costs of skill acquisition, her expected *gross* income from acquiring the skill is $w(1 - F_q(s_{ij}))$, where the expectation is over the level of signal this worker expects to send. Similarly, her expected income from not acquiring the skill is $w(1 - F_u(s_{ij}))$. Letting $\beta(s_{ij})$ denote the *increase* in expected gross income from skill acquisition, we can write $\beta(s_{ij}) \equiv w(1 - F_q(s_{ij})) - w(1 - F_u(s_{ij})) = w[F_u(s_{ij}) - F_q(s_{ij})]$. It is straightforward to check that $\beta(0) = \beta(1) = 0$, so that $G(\beta(0)) = G(\beta(1)) = 0$. Further, given MLRP, $G(\beta(s_{ij})) \geq 0$ and increasing if and only if $\phi(s_{ij}) > 1$. This in turn implies that $G(\beta(s_{ij}))$ is single-peaked, attaining a maximum value at some $\tilde{s} > 0$, so that $G(\beta(s_{ij}))$ is increasing if and only if $s_{ij} < \tilde{s}$.

First consider rich workers. A rich worker with type c and identity i acquires the skill if and only if $w(1 - F_q(s_{iR})) - c \geq w(1 - F_u(s_{iR}))$, i.e. the cost of skill acquisition

$$c \leq \beta(s_{iR}) \equiv w(F_u(s_{iR}) - F_q(s_{iR})), \tag{7}$$

the expected gain from doing so. Recalling that c has a distribution $G(c)$, the proportion of category iR workers getting educated, i.e. π_{iR}, equals the proportion of workers with cost less than $\beta(s_{iR})$, so that

$$\pi_{iR} = G(\beta(s_{iR})), \ i = \{B, W\}. \tag{8}$$

Given that $G(\beta(s))$ is single peaked at $\tilde{s}$ and increasing if and only if $s < \tilde{s}$, the graph of $\pi_{iR}(s_{iR})$ in the $s - \pi$ space, call it WW_R (where WW is a mnemonic for *w*orkers), is inversely U-shaped.

Recalling that for a poor worker of type c the cost of skill acquisition is cm, we can mimic the preceding argument to write (denoting $\hat{\beta}(s) \equiv \frac{\beta(s)}{m}$)

$$\pi_{iP} = G(\hat{\beta}(s_{iP})), \ i = \{B, W\}. \tag{9}$$

Given that $G(\hat{\beta}(s))$ is single peaked at $\tilde{s}$ and increasing if and only if $s < \tilde{s}$, the graph of $\pi_{iP}(s_{iP})$ in the $s - \pi$ space, call it WW_P, is inversely U-shaped with a peak at $\tilde{s}$ (see Figure 1).

We next define the notion of an equilibrium in the absence of affirmative action.

An equilibrium for a group $\{i, j\}$, $i \in \{W, B\}$, $j \in \{P, R\}$, is a pair $(\bar{s}_{ij}, \bar{\pi}_{ij})$ such that (a) given the cutoff $\bar{s}_{ij}$, the proportion of group $\{i, j\}$ workers acquiring the skill is $\bar{\pi}_{ij}$, and (b) given the level of skill acquisition $\bar{\pi}_{ij}$, a cut-off of $\bar{s}_{ij}$ maximises firm profits, i.e. $(\bar{s}_{iR}, \bar{\pi}_{iR})$ satisfies (8), and $(\bar{s}_{iP}, \bar{\pi}_{iP})$ satisfies (9).

An *equilibrium* for this economy is a configuration $< \bar{s}_{ij}, \bar{\pi}_{ij} >$, $i \in \{W, B\}$, $j \in \{P, R\}$, such that $(\bar{s}_{ij}, \bar{\pi}_{ij})$ is an equilibrium for the group $\{i, j\}$.

We focus on equilibria that are *locally stable* in that the absolute value of the slope of EE exceeds that of WW_j, $j = \{P, R\}$, at the equilibrium levels of π and s.[14] Next note that for any group $\{i, j\}$ there always exists a trivial equilibrium such that $\pi_{ij} = 0$ and $s_{ij} = 1$, so that the firm believes that no agent in this group is qualified, and based on this belief assigns everyone from this group to task 2 making this belief self-fulfilling. We shall examine equilibria that are not trivial in that there is at least one group that acquires a positive level of skills. Finally, given that the fundamental problem of interest is the stereo-typing of black workers in the absence of any affirmative action, we shall examine equilibria $< \bar{s}_{ij}, \bar{\pi}_{ij} >$ where blacks are held to standards that are higher than that of the white-workers, i.e. $\bar{s}_{Bj} \geq \bar{s}_{Wj}$, $j = \{P, R\}$. We shall focus on equilibria that satisfy all these properties, referring to them simply as equilibria for ease of exposition.

We next impose an assumption that ensures that there exist group equilibria with positive levels of skill acquisition for all workers. In fact, we shall impose a stronger condition, assumption 2 below, that will help ensure the existence of equilibria even when affirmative action constraints are imposed.

Assumption 2. *Defining* $\tilde{r}$ *as solving* $G(\beta(\tilde{s})) = \frac{\phi(\tilde{s})}{r+\phi(\tilde{s})}$, *and* $\tilde{z}$ *as solving* $G(\hat{\beta}(\tilde{s})) = \frac{\phi(\tilde{s})}{z+\phi(\tilde{s})}$:

(a) $\frac{(\lambda_R)x_q-(\lambda_P)x_u-w}{x_u+w} > \tilde{r}$.

(b) $\frac{(\lambda_W)x_q-(\lambda_B)x_u-w}{x_u+w} > \tilde{r}$.

(c) $\frac{x_q-w}{x_u+w} > \tilde{z}$.

Assumption 2 ensures there is a unique $s' < \tilde{s}$ (respectively $s'' < \tilde{s}$) such that EE and WW_R (respectively WW_P) intersect (see Figure 1).[15]

[14] This follows from a dynamic adjustment protocol given by $\pi^{t+1} = G(\beta(s(\pi^t)))$, $t = 0, 1, 2, \ldots$ for white workers and $\pi^{t+1} = G(\hat{\beta}(s(\pi^t)))$, $t = 0, 1, 2, \ldots$ for black workers.

[15] Given that WW_P lies below WW_R for all $0 < s < 1$, it is sufficient to show that there exists $s'' < \tilde{s}$ such that EE and WW_P intersect, since it implies that there exists $s' < \tilde{s}$ such that EE and WW_R intersect. From (9), note that along the WW_P curve we have that $\pi_{WW_P}(\tilde{s}) = G(\hat{\beta}(\tilde{s}))$, whereas from (6) we have that $\pi_{EE}(\tilde{s}, z) = \frac{\phi(\tilde{s})}{z+\phi(\tilde{s})}$ along the EE curve. Thus for all $z > \tilde{z}$, we have that $\pi_{WW_P}(\tilde{s}) > \frac{\phi(\tilde{s})}{z+\phi(\tilde{s})} = \pi_{EE}(\tilde{s}, z)$ (from assumption 2(b)).

We are interested in economies (and equilibria) where identity-based stereo-typing is a serious problem, in that whenever stereo-typing exists at a certain income level, black workers at that income level are held to the highest possible standard (i.e. cutoff), while white workers at the same income level are held to the lowest possible standard (in a sense formalized later). We shall examine three different benchmark equilibria that capture different assumptions regarding the role of income mobility in alleviating such stereo-typing, as well as how debilitating poverty is to skill acquisition.

The first equilibrium, denoted $\mathcal{E}^{\mathsf{P}}$ (where P stands for poverty), captures the possibility that for the poor, poverty is so debilitating that getting skilled is not really an option so that all poor workers, both black and white, are subject to the same high standard. For the not-so-poor however, stereo-typing is still a fact of life. The second equilibrium, denoted $\mathcal{E}^{\mathsf{ST}}$ (where ST stands for stereo-typing), is one where black workers are stereo-typed irrespective of their income level, thus capturing the idea that stereo-typing is so deep-seated that higher income is no guarantee of escaping it. The third equilibrium, denoted $\mathcal{E}^{\mathrm{IM}}$ (where IM stands for income mobility), captures the opposite idea that upward economic mobility may allow rich black workers to escape stereo-typing (see Figure 1).[16]

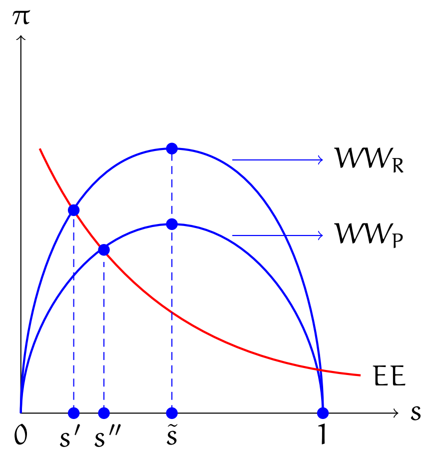
We take the position that depending on time and place, any of these equilibria may be salient. Moreover, with societal and economic changes, the economy may very well transition across these equilibria.

$\mathcal{E}^{\mathsf{P}}$: An equilibrium $< \bar{s}_{ij}, \bar{\pi}_{ij} >$ is said to be $\mathcal{E}^{\mathsf{P}}$ if (a) no poor workers, black or white, are assigned to task 1, i.e. $\bar{s}_{BP} = \bar{s}_{WP} = 1$, and (b) rich white workers are not only subject to a lower standard relative to the rich black workers, i.e. $\bar{s}_{BR} > \bar{s}_{WR}$, moreover there exists no other group-equilibrium for the rich white workers where they are subject to a lower standard, as well as no other group-equilibrium for the rich black workers where they are subject to a higher standard.

$\mathcal{E}^{\mathsf{ST}}$: An equilibrium $< \bar{s}_{ij}, \bar{\pi}_{ij} >$ is said to be $\mathcal{E}^{\mathsf{ST}}$ if (a) black workers, either poor or rich, are never assigned to task 1, i.e. $\bar{s}_{BP} = \bar{s}_{BR} = 1$, (b) white workers are not only subjected to a lower standard vis-á-vis black workers with the same income level, i.e. $\bar{s}_{WP}, \bar{s}_{WR} < 1$, moreover, there exists no other group equilibrium for the white workers where they are held to a lower standard.

$\mathcal{E}^{\mathrm{IM}}$: An equilibrium $< \bar{s}_{ij}, \bar{\pi}_{ij} >$ is said to be $\mathcal{E}^{\mathrm{IM}}$ if (a) all rich workers, black or white, are subject to the same standards, i.e. $\bar{s}_{BR} = \bar{s}_{WR}$, and there exists no other group-equilibrium for the rich workers where they are subject to a lower standard, (b) poor black workers are never assigned to task 1, i.e. $\bar{s}_{BP} = 1$, and (c) poor white workers are not only subject to a lower standard relative to the poor black workers, i.e. $\bar{s}_{WP} < 1$, moreover there exists no other group-equilibrium for the poor white workers where they are subject to a lower standard.

[16] This idea underpins the fact that in some countries, e.g. India, affirmative action policy excludes the rich among the intended beneficiary group.



$\mathcal{E}^{P}$: $\bar{s}_{WR} = s'$, $\bar{s}_{BR} = \bar{s}_{BP} = \bar{s}_{WP} = 1$.

$\mathcal{E}^{ST}$: $\bar{s}_{WR} = s'$, $\bar{s}_{WP} = s''$, $\bar{s}_{BR} = \bar{s}_{BP} = 1$.

$\mathcal{E}^{IM}$: $\bar{s}_{WR} = \bar{s}_{BR} = s', \bar{s}_{WP} = s''$, $\bar{s}_{BP} = 1$.

Figure 1: *Equilibria in the Absence of Affirmative Action*

3.1 Welfare Comparisons

Before turning to an analysis of equilibria with affirmative action, we derive some results that will be required while comparing the three different regimes - no affirmative action, class-based affirmative action, and identity-based affirmative action. In particular we show that, comparing across equilibria and even different affirmative action regimes, any given group of workers is better off under a particular equilibrium relative to another iff they face a lower cutoff under this equilibrium relative to the other equilibrium. Moreover, firm profit is higher under a particular equilibrium relative to another whenever all groups acquire a higher level of skills under this equilibrium. The proof essentially follows from a revealed preference argument.

Proposition 1. *Consider two distinct equilibrium outcomes,* $< s'_{ij}, \pi'_{ij}, \rho'_{ij} >$*, and* $< s''_{ij}, \pi''_{ij}, \rho''_{ij} >$*.*

(A) Every individual in group {i,j} prefers $< s'_{ij}, \pi'_{ij}, \rho'_{ij} >$ *to* $< s''_{ij}, \pi''_{ij}, \rho''_{ij} >$ *if and only if they face a relatively lower signal cutoff under* $< s'_{ij}, \pi'_{ij}, \rho'_{ij} >$*, i.e.* $s'_{ij} < s''_{ij}$*.*

(B) If all groups acquire a higher skill level under $< s'_{ij}, \pi'_{ij}, \rho'_{ij} >$*, i.e.* $\pi'_{ij} \geq \pi''_{ij}$*,* $\forall\{i, j\}$*, then the firm's profits are higher under* $< s'_{ij}, \pi'_{ij}, \rho'_{ij} >$*.*

Proof. For ease of exposition, let us denote $< s'_{ij}, \pi'_{ij}, \rho'_{ij} >$ by E1, and $< s''_{ij}, \pi''_{ij}, \rho''_{ij} >$ by E2.

(A) We divide the individuals in group {i,j} into two categories.

(a) First consider group {i,j} individuals who take the same decision regarding skill acquisition under both equilibria, i.e. they either acquire the skill under both E1 and E2, or refuse to do so under both equilibria. Given that their level of skill is the same under both equilibria, they prefer E1 over E2 since, given that $s'_{ij} < s''_{ij}$, they have a greater chance of being assigned to task 1 under E1.

(b) Next consider individuals whose skill acquisition decision changes across the two equilibria. Let $u_{ij}(c, x, s, Ek)$ denote the utility of a group {i,j} individual with cost c under equilibrium Ek facing a cutoff signal of s, and taking a skill acquisition decision $x \in \{Y, N\}$, with Y (resp. N) denoting that she decides to acquire (resp. not acquire) the skill.

(i) First, consider individuals who acquire the skill under E2, but not under E1. Then

$$\begin{aligned} u_{ij}(c, N, s'_{ij}, E1) &\geq u_{ij}(c, Y, s'_{ij}, E1) &(10)\\ &> u_{ij}(c, Y, s''_{ij}, E2), &(11) \end{aligned}$$

where the first inequality follows from a revealed preference argument and the second inequality from the fact that $s'_{ij} < s''_{ij}$. Thus this individual is strictly better off under E1.

(ii) Next, consider individuals who acquire the skill under E1, but not under E2. Then

$$\begin{aligned} u_{ij}(c, Y, s'_{ij}, E1) &\geq u_{ij}(c, N, s'_{ij}, E1) &(12)\\ &> u_{ij}(c, N, s''_{ij}, E2). &(13) \end{aligned}$$

Thus all group {i,j} workers are strictly better off under E1 relative to E2.

(B) The firm's profits under $< s'_{ij}, \pi'_{ij}, \rho'_{ij} >$

$$\begin{aligned} \sum_{i=W,B} \sum_{j=R,P} \lambda_{ij} P(s'_{ij}, \pi'_{ij}) &\geq \sum_{i=W,B} \sum_{j=R,P} \lambda_{ij} P(s''_{ij}, \pi'_{ij}) \\ &\geq \sum_{i=W,B} \sum_{j=R,P} \lambda_{ij} P(s''_{ij}, \pi''_{ij}), &(14) \end{aligned}$$

which equals the profits under $< s''_{ij}, \pi''_{ij}, \rho''_{ij} >$. Note that the first inequality follows from a revealed preference argument (since firms set s'_{ij} optimally under the equilibrium $< s'_{ij}, \pi'_{ij}, \rho'_{ij} >$), and the second inequality follows from the fact that $P(s_{ij}, \pi_{ij})$ is increasing in π_{ij}. □

4 Class-based affirmative action

In this section we examine the effect of a class-based affirmative action policy. We show that there exists a CAA equilibrium that is colour-blind, i.e. all poor workers face the same cut-offs irrespective of their identity, as do the rich. Further, in any such equilibrium the poor face lower cutoffs vis-á-vis the rich.

We begin by analysing the decision problems facing the firm.

The firm's decision: Let the firm hold the belief that a fraction π_{ij} of the workers in the group {i,j} got skilled. Thus the constrained optimization problem facing the firm is to choose cutoffs for each category of worker, assigning them to task 1 if their signal exceeds these cutoffs, subject to the CAA constraint. Thus the Lagrange can be written as:

$$\max_{s'_{BP}, s'_{BR}, s'_{WP}, s'_{WR}} \sum_{i=W,B} \sum_{j=R,P} \lambda_{ij} P(s'_{ij}, \pi_{ij}) + \gamma \sum_{i=B,W} \left[\frac{\lambda_{iP}}{\lambda_P} \rho_{iP}(s'_{iP}, \pi_{iP}) - \frac{\lambda_{iR}}{\lambda_R} \rho_{iR}(s'_{iR}, \pi_{iR}) \right], \quad (15)$$

where γ is the Lagrange multiplier on the CAA constraint. The first order condition with respect to s'_{ij} generates the cutoff s_{ij} as a function of π_{ij}. Denoting this function by $EE_{ij}(\gamma)$, we have:

$$
\begin{aligned}
\frac{x_q - w + \gamma/\lambda_P}{x_u + w - \gamma/\lambda_P} &= \tfrac{1-\pi_{BP}}{\pi_{BP}}\phi(s_{BP}) \quad : \; EE_{BP}(\gamma), && (16)\\
\frac{x_q - w - \gamma/\lambda_R}{x_u + w + \gamma/\lambda_R} &= \tfrac{1-\pi_{BR}}{\pi_{BR}}\phi(s_{BR}) \quad : \; EE_{BR}(\gamma), && (17)\\
\frac{x_q - w + \gamma/\lambda_P}{x_u + w - \gamma/\lambda_P} &= \tfrac{1-\pi_{WP}}{\pi_{WP}}\phi(s_{WP}) \quad : \; EE_{WP}(\gamma), && (18)\\
\frac{x_q - w - \gamma/\lambda_R}{x_u + w + \gamma/\lambda_R} &= \tfrac{1-\pi_{WR}}{\pi_{WR}}\phi(s_{WR}) \quad : \; EE_{WR}(\gamma). && (19)
\end{aligned}
$$

These conditions are intuitive, showing that as a result of affirmative action, the firm acts as if it has to pay a tax of $\frac{\gamma}{\lambda_R}$ on each rich worker assigned to task one, while receiving subsidies of $\frac{\gamma}{\lambda_P}$ on each poor worker it assigns to task one. Further, the EE curves for rich black and rich white workers, i.e. EE_{BR} and EE_{WR} coincide, as do the ones for poor whites and poor blacks, i.e. EE_{WP} and EE_{BP}. Thus, for any $\gamma > 0$, we have two EE curves, which we can denote by $EE_R(\gamma)$ and $EE_P(\gamma)$, with the former lying to the right of the latter in $s - \pi$ space. Further, from MLRP it follows that s_{ij} is decreasing in π_{ij}, so that the EE_j curves are negatively sloped in the $s - \pi$ space. Note that as γ tends to zero, $EE_R(\gamma)$ and $EE_P(\gamma)$ coincide with EE.

Note that the decision problem of the workers can still be summarized by WW_R for the rich, and by WW_P for the poor workers.

We then define the functions

$$
\begin{aligned}
\hat{\rho}(s) &\equiv \rho(s, G(\beta(s)) = G(\beta(s))[1 - F_q(s)] + (1 - G(\beta(s)))[1 - F_u(s)], && (20)\\
\text{and } \hat{\hat{\rho}}(s) &\equiv \rho(s, G(\hat{\beta}(s)) = G(\hat{\beta}(s))[1 - F_q(s)] + (1 - G(\hat{\beta}(s)))[1 - F_u(s)], && (21)
\end{aligned}
$$

that will play an important role in the analysis. For $i \in \{B, W\}$, $\hat{\rho}(s_i)$ denotes the fraction of rich workers with identity i that will be assigned to the skilled task when the firm adopts a cutoff of s_i for rich workers with identity i, and the workers respond optimally to this cutoff so that a proportion $G(\beta(s_i))$ of workers in this group invest in skill acquisition. Similarly, $\hat{\hat{\rho}}(s_i)$ denotes the fraction of poor workers with idntity i, $i = \{B, W\}$, that will be assigned to the skilled task when the firm has a cutoff of s_i for poor workers with identity i, and the workers respond optimally to this cutoff. Note that $\hat{\rho}(0) = \hat{\hat{\rho}}(0) = 1$ and $\hat{\rho}(1) = \hat{\hat{\rho}}(1) = 0$. Moreover, given that $\hat{\beta}(s) = \frac{\beta(s)}{m} < \beta(s)$, it is clear that $\hat{\hat{\rho}}(s) < \hat{\rho}(s)$, $\forall s > 0$.

We then impose some restrictions on the $\hat{\rho}(s)$ and $\hat{\hat{\rho}}(s)$ functions.

Assumption 3. $\hat{\rho}(s)$ *and* $\hat{\hat{\rho}}(s)$ *are negatively sloped.*

Assumption 3 ensures that, even taking into account the fact that the workers respond optimally to any changes in the cutoff, an increase in the cutoff signal worsens the chances that an worker emitting a particular signal is assigned to the high wage task.[17]

[17]In terms of model primitives, that $\hat{\rho}(s)$ is negatively sloped can equivalently be written as $\frac{f_u(s)-f_q(s)}{f_u(s)} <$

An outcome $< s^{**}_{ij}, \pi^{**}_{ij}, \rho^{**}_{ij} >$, where $i = B, W$, $j = R, P$, constitutes an *equilibrium with CAA* if and only if, (a) given that all workers respond optimally to the cutoff $< s^{**}_{ij} >$, the proportion of workers of the relevant category acquiring the skill is π^{**}_{ij}, (b) given the vector of skill acquisition $(\pi^{**}_{BP}, \pi^{**}_{BR}, \pi^{**}_{WP}, \pi^{**}_{WR})$, the cut-off vector of $(s^{**}_{BP}, s^{**}_{BR}, s^{**}_{WP}, s^{**}_{WR})$ maximises firm profits subject to the class-based affirmative action constraint (2), and (c) $\rho^{**}_{ij} = \rho_{ij}(s^{**}_{ij}, \pi^{**}_{ij})$.

Let γ^{**} denote the Lagrange multiplier associated with this equilibrium. As in the baseline model without affirmative action, we shall restrict attention to equilibria that (a) are locally stable, so that for all i, $EE_j(\gamma^{**})$ intersects WW_j from above, (b) does not involve all groups acquiring zero skills. We shall be interested in what we call

Colour-blind equilibria: A CAA equilibrium $< s^{**}_{ij}, \pi^{**}_{ij}, \rho^{**}_{ij} >$ is said to be colour-blind iff $s^{**}_{BP} = s^{**}_{WP}$ and $s^{**}_{BR} = s^{**}_{WR}$.

In Proposition 2 below we argue that a colour-blind equilibrium exists (see Figure 2). Furthermore, this equilibrium is unique in the class of non-trivial colour-blind equilibria.

Proposition 2. *Let assumptions 1, 2, and 3 hold.*

(a) There exists a locally stable CAA equilibrium that has a positive Lagrange multiplier, and is colour-blind in that $s^{**}_{BP} = s^{**}_{WP}$ *and* $s^{**}_{BR} = s^{**}_{WR}$. *Moreover, there is exactly one such non-trivial colour-blind equilibrium (i.e. the skill-level of at least one group is positive).*

(b) Consider any colour-blind CAA with a positive Lagrange multiplier. Under any such equilibrium,

- *(i) all workers have equal chances of being assigned to task 1, i.e.* $\rho^{**}_{BP} = \rho^{**}_{WP} = \rho^{**}_{BR} = \rho^{**}_{WR}$,
- *(ii) the poor face a lower cut-off vis-á-vis the rich, i.e.* $s^{**}_{BP} = s^{**}_{WP} < s^{**}_{BR} = s^{**}_{WR}$, *and*
- *(iii)* $s^{**}_{BR} = s^{**}_{WR} < 1$.

Proof. (a) Define $\hat{s}_{ij}$ as solving

$$\hat{s}_{ij}(\gamma) = \{\min s_{ij} | s_{ij} \text{ lies at the intersection of } \mathrm{EE}_j(\gamma) \text{ and } \mathrm{WW}_j\}.$$

Note that $\hat{s}_{BR}(0) = \hat{s}_{WR}(0) = s'$ and $\hat{s}_{BP}(0) = \hat{s}_{WP}(0) = s''$. Further, we have that $\hat{s}_{iR}(0) = s' < \hat{s}_{iP}(0) = s''$ (since $m > 1$), and since $r > \tilde{r}$, we have that $s' < s''$ (see Figure 1). Thus from assumption 3 and the fact that $s' < s''$, and that $\hat{\rho}(s) > \hat{\hat{\rho}}(s)\ \forall s > 0$, one has that

$$\hat{\rho}(s') > \hat{\hat{\rho}}(s''). \tag{22}$$

$\frac{1}{G(\beta(s))+g(\beta(s))\beta(s)}$, $\forall s$, and that $\hat{\hat{\rho}}(s)$ is negatively sloped can equivalently be written as $\frac{f_u(s)-f_q(s)}{f_u(s)} < \frac{1}{G(\hat{\beta}(s))+g(\hat{\beta}(s))\hat{\beta}(s)}$, $\forall s$. If, for example, (a) c is uniformly distributed over $[0,1]$, so that $g(.) = 1$, (b) $f_u(s) = 2$ for all $s \in [0, 1/2]$ and $f_u(s) = 0$, $\forall s \in (1/2, 1]$, (c) $f_q(s) = 1$ for all $s \in [0, 1]$, and (d) $w < 1$, then it is straightforward to check that the conditions both hold.

We have to show that there exists $\gamma^{**} > 0$, and $\hat{s}_{ij}(\gamma^{**})$ such that (i) $\hat{s}_{BP}(\gamma^{**}) = \hat{s}_{WP}(\gamma^{**})$, (ii) $\hat{s}_{BR}(\gamma^{**}) = \hat{s}_{WR}(\gamma^{**})$ and (iii) the class-based affirmative action constraint (2) is satisfied by $\hat{s}_{ij}(\gamma^{**})$. We shall use the intermediate value theorem to establish the existence of such a γ^{**}. To that end we first define

$$\begin{aligned} F(\gamma) &= [\frac{\lambda_{BP} + \lambda_{WP}}{\lambda_P}]\hat{\hat{\rho}}(\hat{s}_{BP}(\gamma)) - [\frac{\lambda_{BR} + \lambda_{WR}}{\lambda_R}]\hat{\rho}(\hat{s}_{BR}(\gamma)) \\ &= [\hat{\hat{\rho}}(\hat{s}_{BP}(\gamma)) - \hat{\rho}(\hat{s}_{BR}(\gamma))], \end{aligned} \tag{23}$$

where the last equality follows since $\frac{\lambda_{BP}+\lambda_{WP}}{\lambda_P} = \frac{\lambda_{BR}+\lambda_{WR}}{\lambda_R} = 1$.

We then note that from (22), we have that

$$F(0) = [\hat{\hat{\rho}}(\hat{s}_{BP}(0)) - \hat{\rho}(\hat{s}_{BR}(0))] = [\hat{\hat{\rho}}(s'') - \hat{\rho}(s')] < 0.$$

Next let γ increase from zero. Observe that $\hat{s}_{ij}(\gamma)$ is continuous, where the continuity of $\hat{s}_{iR}(\gamma)$ follows since from assumption 2(a), $EE_R(\gamma)$ and WW_R always has an intersection for all $\gamma \leq \lambda_P(x_u + w)$.[18] Moreover, $\hat{s}_{BP}(\gamma)$ and $\hat{s}_{WP}(\gamma)$ are decreasing, and $\hat{s}_{WR}(\gamma)$ and $\hat{s}_{BR}(\gamma)$ are increasing in γ.

Further, as $\gamma \to \lambda_P(x_u + w)$, $\hat{s}_{iP}(\gamma)$ goes to zero (from (16) and (18)),[19] so that from (21) $\hat{\hat{\rho}}(s_{iP})$ goes to 1,[20] and consequently $\lim_{\gamma\to\lambda_P(x_u+w)} \hat{\hat{\rho}}(\hat{s}_{BP}(\gamma)) = 1$. Whereas $\hat{s}_{iR}(\gamma) > 0$, so that $\hat{\rho}(\hat{s}_{iR}(\gamma)) < 1$, and consequently $F(\gamma)|_{\gamma\to\lambda_P(x_u+w)} > 1 - 1 = 0$. Thus there exists $\gamma^{**}(m) > 0$, such that $\hat{s}_{ij}(\gamma^{**})$ satisfies CAA.

That the equilibrium is locally stable follows from the fact that, by construction, at the equilibrium the absolute value of the slope of EE_R (respectively EE_P) exceeds that of WW_R (respectively WW_P). This follows since EE_j are negatively sloped and WW_j are positively sloped for $s < \tilde{s}$.

We then argue that there is exactly one such non-trivial colour-blind equilibrium. Suppose to the contrary there exists another colour-blind CAA equilibrium with Lagrange multiplier γ', where say $\gamma' < \gamma^{**}$. This implies that the cut-off signals facing the poor would be larger, and that facing the rich would be smaller under this equilibrium relative to that under γ^{**}. Given Assumption 3, this in turn ensures that the poor would be assigned at a lower rate to task 1, and the rich would be assigned to task 1 at a higher rate under this equilibrium relative to the equilibrium under γ^{**}. However, this implies that the CAA would be violated under this equilibrium, given that it was being satisfied under γ^{**}. A similar argument

[18]From (8), note that along the WW_R curve we have that $\pi_{WW_R}(\tilde{s}) = G(\beta(\tilde{s}))$, whereas from (17) (and (19)) we have that $\pi_{EE_R}(\tilde{s}, r, \gamma = \lambda_P(x_u + w)) = \frac{\phi(\tilde{s})}{r+\phi(\tilde{s})}$ along the EE_R curve. Thus for all $r > \tilde{r}$, we have that $\pi_{WW_R}(\tilde{s}) > \frac{\phi(\tilde{s})}{r+\phi(\tilde{s})} = \pi_{EE_R}(\tilde{s}, r, \gamma = \lambda_P(x_u + w))$ (from assumption 2(a)).

[19]Note that from (16), we have that $\lim_{\gamma\to\lambda_P(x_u+w)} \phi(s_{BP}) = \infty$. Given that $\phi(s)$ is decreasing, and that $\lim_{s\to 0} \phi(s) = \infty$ (Assumption 1), we have that $\lim_{\gamma\to\lambda_P(x_u+w)} s_{BP} = 0$. A similar argument holds for s_{WP}.

[20]Note that, $F_q(0) = F_u(0) = 0$, so that

$$\begin{aligned} \lim_{s\to 0} \hat{\hat{\rho}}(s) &= \lim_{s\to 0} [G(\hat{\beta}(s))[1 - F_q(s)] + (1 - G(\hat{\beta}(s)))[1 - F_u(s)]] \\ &= \lim_{s\to 0} [G(\hat{\beta}(s)) + (1 - G(\hat{\beta}(s))] = 1. \end{aligned}$$

applies for $\gamma' > \gamma^{**}$.

(b)(i) Consider a colour-blind equilibrium so that $s^{**}_{BP} = s^{**}_{WP}$ and $s^{**}_{BR} = s^{**}_{WR}$. Given Assumption 3, this ensures that $\rho^{**}_{BP} = \rho^{**}_{WP}$ and $\rho^{**}_{BR} = \rho^{**}_{WR}$. Next recall from the class-based affirmative action constraint (2) that,

$$\frac{\lambda_{BP}}{\lambda_P}\rho^{**}_{BP} + \frac{\lambda_{WP}}{\lambda_P}\rho^{**}_{WP} = \frac{\lambda_{BR}}{\lambda_R}\rho^{**}_{BR} + \frac{\lambda_{WR}}{\lambda_R}\rho^{**}_{WR}.$$

The preceding equation, along with the fact that $\rho^{**}_{BP} = \rho^{**}_{WP}$ and $\rho^{**}_{BR} = \rho^{**}_{WR}$, ensures that

$$\rho^{**}_{WP} = [\frac{\lambda_{BP} + \lambda_{WP}}{\lambda_P}]\rho^{**}_{WP} \quad = \quad [\frac{\lambda_{BR} + \lambda_{WR}}{\lambda_R}]\rho^{**}_{WR} = \rho^{**}_{WR}, \tag{24}$$

since $\frac{\lambda_{BP}+\lambda_{WP}}{\lambda_P} = \frac{\lambda_{BR}+\lambda_{WR}}{\lambda_R} = 1$.

(b)(ii) Suppose to the contrary that $s^{**}_{BP} \geq s^{**}_{BR}$. Given Assumption 3, this ensures that

$$\hat{\rho}(s^{**}_{BP}) \leq \hat{\rho}(s^{**}_{BR}).$$

Given the fact that $\hat{\hat{\rho}}(s) < \hat{\rho}(s)\ \forall s > 0$, the preceding inequality implies that

$$\hat{\hat{\rho}}(s^{**}_{BP}) < \hat{\rho}(s^{**}_{BP}) \leq \hat{\rho}(s^{**}_{BR}).$$

This however contradicts the fact that from Proposition 2(b)(i), $\hat{\hat{\rho}}(s^{**}_{BP}) = \hat{\rho}(s^{**}_{BR})$.

(b)(iii) Suppose to the contrary that $s^{**}_{BR} = s^{**}_{WR} = 1$. This implies that $\rho^{**}_{BR} = \rho^{**}_{WR} = 0$, so that the CAA constraint can only be sustained if the equilibrium is a trivial one. □

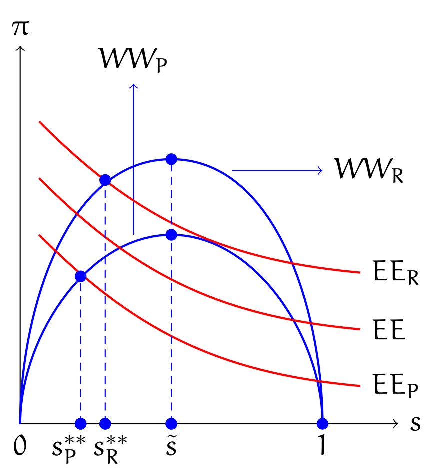


Figure 2: *Color-blind equilibrium under CAA:* $s^{**}_{BP} = s^{**}_{WP} \equiv s^{**}_P$, $s^{**}_{WR} = s^{**}_{BR} \equiv s^{**}_R$

4.1 Comparing the colour-blind equilibrium under CAA with those in the benchmark model

We then examine how the equilibria under CAA compares with those in the benchmark model with no affirmative action, namely $\mathcal{E}^{\mathsf{P}}$, $\mathcal{E}^{\mathsf{ST}}$ and $\mathcal{E}^{\mathsf{IM}}$. While, as is to be expected, the result depends on which one of the equilibria prevails in the absence of affirmative action, some broad patterns can be discerned.

First, consider the effect on poor workers. We find that irrespective of the prevailing equilibrium, all poor workers are better off under CAA. Whether skill acquisition improves or not is however sensitive to the nature of the equilibrium that prevails in the absence of any affirmtaive action. While skill acquisition by *all* poor workers increases relative to $\mathcal{E}^{\mathsf{P}}$, poor white workers acquire less skill under CAA relative to both $\mathcal{E}^{\mathsf{ST}}$ and $\mathcal{E}^{\mathsf{IM}}$.

Turning to the effect of CAA on the black workers, we find that while all black workers gain relative to $\mathcal{E}^{\mathsf{P}}$ and $\mathcal{E}^{\mathsf{ST}}$, relative to $\mathcal{E}^{\mathsf{IM}}$ rich black workers are actually worse off under CAA. Thus rich black workers who are not subject to stereo-typing lose out because of CAA. As far as skill acquisition is concerned, while all black workers acquire a greater amount of skill under CAA relative to both $\mathcal{E}^{\mathsf{P}}$ and $\mathcal{E}^{\mathsf{ST}}$, relative to $\mathcal{E}^{\mathsf{IM}}$ poor black workers acquire a greater amount of skill under CAA.

Proposition 3. *Let assumptions 1, 2, and 3 hold.*

(a) Compare the $\mathcal{E}^{\mathsf{P}}$ equilibrium with the colour-blind equilibrium under CAA. Under CAA, while rich white workers are worse off, all other workers are better off. Further, skill acquisition by all workers are higher. The firm also has a higher profit.

(b) Compare the $\mathcal{E}^{\mathsf{ST}}$ equilibrium with the colour-blind equilibrium under CAA. Under CAA, all workers, other than the rich white workers are better off. Further, skill acquisition by poor white workers decreases, whereas skill acquisition by black workers, as well as the rich white workers increases.

(c) Compare the $\mathcal{E}^{\mathsf{IM}}$ equilibrium with the colour-blind equilibrium under CAA. Under CAA, all poor workers are better off, whereas all rich workers are worse off. Further, skill acquisition by all black workers, as well as rich white workers increases, whereas skill acquisition by poor white workers decreases.

5 Identity based affirmative action

We next characterise equilibria under identity based affirmative action. We find that under IAA there exists an equilibrium where black workers are held to lower standards relative to white workers. Thus this equilibrium exhibits patronization as identified by Coate and Loury (1993) (we shortly provide a formal definition of patronization). Having characterized this equilibrium, we then compare the outcome under the two forms of affirmative action in the next section.

We examine a society where blacks are poorer on the average vis-á-vis the whites.

Assumption 4. $\mu_{\mathsf{BP}} > \mu_{\mathsf{WP}}$.

Given that the poor find it more costly to acquire the relevant skill set, assumption 4 implies that on the average black workers find it costlier to get skilled.

The firm's decisions: Let the firm hold the belief that a fraction π_{ij} of the workers in the group {i,j} got skilled. Facing a worker with identity i and class j, say, who emits a certain signal, the constrained optimization problem facing the firm is to choose cutoffs for each category of worker, assigning them to task 1 if their signal exceeds these cutoffs, subject to the IAA constraint:

$$\max_{s'_{BP},s'_{BR},s'_{WP},s'_{WR}} \sum_{i=W,B} \sum_{j=R,P} \lambda_{ij} P(s'_{ij}, \pi_{ij}) + \gamma \sum_{j=R,P} [\mu_{Bj}\rho_{Bj}(s'_{Bj}, \pi_{Bj}) - \mu_{Wj}\rho_{Wj}(s'_{Wj}, \pi_{Wj})], \tag{25}$$

where γ is the Lagrange multiplier on the IAA constraint. The first order condition with respect to s'_{ij} generates the cutoff s_{ij} as a function of π_{ij}. Denoting this function by $EE_{ij}(\gamma)$, we have

$$\frac{x_q - w + \gamma\mu_{BP}/\lambda_{BP}}{x_u + w - \gamma\mu_{BP}/\lambda_{BP}} = \frac{x_q - w + \gamma/\lambda_B}{x_u + w - \gamma/\lambda_B} = \tfrac{1-\pi_{BP}}{\pi_{BP}}\phi(s_{BP}) \quad : \; EE_{BP}(\gamma), \tag{26}$$

$$\frac{x_q - w + \gamma\mu_{BR}/\lambda_{BR}}{x_u + w - \gamma\mu_{BR}/\lambda_{BR}} = \frac{x_q - w + \gamma/\lambda_B}{x_u + w - \gamma/\lambda_B} = \tfrac{1-\pi_{BR}}{\pi_{BR}}\phi(s_{BR}) \quad : \; EE_{BR}(\gamma), \tag{27}$$

$$\frac{x_q - w - \gamma\mu_{WP}/\lambda_{WP}}{x_u + w + \gamma\mu_{WP}/\lambda_{WP}} = \frac{x_q - w - \gamma/\lambda_W}{x_u + w + \gamma/\lambda_W} = \tfrac{1-\pi_{WP}}{\pi_{WP}}\phi(s_{WP}) \quad : \; EE_{WP}(\gamma), \tag{28}$$

$$\frac{x_q - w - \gamma\mu_{WR}/\lambda_{WR}}{x_u + w + \gamma\mu_{WR}/\lambda_{WR}} = \frac{x_q - w - \gamma/\lambda_W}{x_u + w + \gamma/\lambda_W} = \tfrac{1-\pi_{WR}}{\pi_{WR}}\phi(s_{WR}) \quad : \; EE_{WR}(\gamma). \tag{29}$$

The EE curves for rich and poor whites coincide, as do the ones for rich and poor blacks. Thus, we have two EE curves, which we can denote by EE_W and EE_B. From MLRP it follows that s_i is decreasing in π_i, where $i \in$ {B,W}. Consequently the EE_i curves are negatively sloped in $s - \pi$ space.

An outcome $< s^*_{ij}, \pi^*_{ij}, \rho^*_{ij} >$, for $i = B, W$, $j = R, P$, constitutes an *equilibrium with IAA* if and only if (a) given that all workers respond optimally to $< s^*_{ij} >$, the proportion of group i and class j workers acquiring the skill is π^*_{ij}, (b) given the vector of skill acquisition $(\pi^*_{BP}, \pi^*_{BR}, \pi^*_{WP}, \pi^*_{WR})$, the cut-off vector $(s^*_{BP}, s^*_{BR}, s^*_{WP}, s^*_{WR})$ maximises firm profits subject to the identity based affirmative action constraint (3), and (c) $\rho^*_{ij} = \rho_{ij}(s^*_{ij}, \pi^*_{ij})$. Formally, $< s^*_{ij}, \pi^*_{ij}, \rho^*_{ij} >$ should satisfy

$$\begin{aligned} \pi_{iR} &= G(\beta(s(\pi_{iR}))), i = B, W, &(30)\\ \pi_{iP} &= G(\hat{\beta}(s(\pi_{iP}))), i = B, W, &(31)\\ \rho_{ij}(s_{ij}, \pi_{ij}) &= \pi_{ij}[1 - F_q(s_{ij})] + (1 - \pi_{ij})[1 - F_u(s_{ij})], &(32)\\ \mu_{WR}\rho_{WR} + \mu_{WP}\rho_{WP} &= \mu_{BR}\rho_{BR} + \mu_{BP}\rho_{BP}. &(33) \end{aligned}$$

Let γ^* denote the Lagrange multiplier associated with this equilibrium. We shall restrict attention to equilibria that are locally stable, so that for all i, $EE_i(\gamma^*)$ intersects WW_j from above. Further, we shall consider equilibria that are non-trivial.

We then define the notion of patronization as in Coate and Loury (1993).

Patronising equilibria: An IAA equilibrium $< s^*_{ij}, \pi^*_{ij}, \rho^*_{ij} >$ is said to be patronizing iff $s^*_{BP} < s^*_{WP}$ and $s^*_{BR} < s^*_{WR}$.

In Proposition 4 below we argue that there exists an equilibrium that exhibits patronization in that black workers of both classes are held to lower standards relative to their white counterparts, i.e. $s^*_{Bj} \leq s^*_{Wj}$, $\forall j \in \{P, R\}$ (see Figure 3). The proof can be found in the appendix.

Proposition 4. *Let assumptions 1, 2, 3 and 4 hold.*

(a) There exists a locally stable IAA equilibrium with a positive Lagrange multiplier where, moreover, black workers are held to lower standards relative to white workers, i.e. $s^*_{Bj} < s^*_{Wj}$, $\forall j \in \{P, R\}$.

(b) Whenever $\mu_{BP} \neq \mu_{WP}$*, there exists no color-blind IAA equilibrium where* $s^*_{BP} = s^*_{WP}$ *and* $s^*_{BR} = s^*_{WR}$.

In our framework this follows since, given that black workers are poorer on the average (Assumption 4), they have higher costs of getting educated on the average as well. Thus firms apply relatively lower cutoffs to black workers of either class while assigning them to task 1, so as to ensure that the IAA constraint holds.[21]

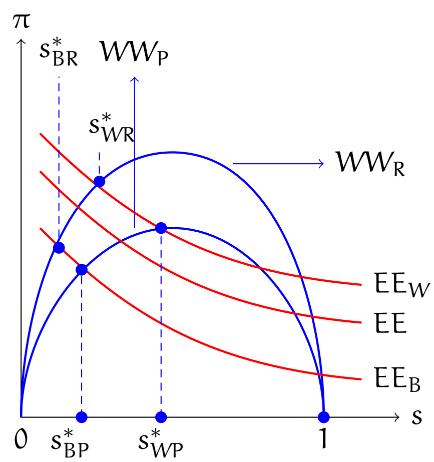


Figure 3: *Patronizing Equilibrium under IAA*

[21] In fact it can be shown that if, instead, the white workers were poorer on the average, then there exists an equilibrium where the white workers, irrespective of their class, would face a lower cutoff relative to their black counter-parts.

5.1 Comparing the patronising equilibrium under IAA with those in the benchmark model

We then examine how the equilibrium under IAA compares with $\mathcal{E}^{\mathsf{P}}$, $\mathcal{E}^{\mathsf{ST}}$ and $\mathcal{E}^{\mathsf{IM}}$.

We find that black workers gain, while rich white workers have a lower utility, under IAA irrespective of which equilibrium prevailed in the absence of affirmative action. As to the skill level, the answer is more nuanced. All black workers get more skilled relative to $\mathcal{E}^{\mathsf{P}}$ and $\mathcal{E}^{\mathsf{ST}}$. Relative to $\mathcal{E}^{\mathsf{IM}}$ however, while rich black workers acquire less skill, poor black workers acquire greater skills.

Turning to poor workers, while poor black workers are necessarily better off, poor white workers are worse off relative to $\mathcal{E}^{\mathsf{ST}}$ and $\mathcal{E}^{\mathsf{IM}}$.

Proposition 5. *Let assumptions 1, 2, 3 and 4 hold.*

(a) Compare the $\mathcal{E}^{\mathsf{P}}$ equilibrium with the patronizing equilibrium in Proposition 4. Under the IAA equilibrium all black workers are better off, whereas rich white workers are worse off. Further, skill acquisition by black workers, as well as rich white workers increases.

(b) Compare the $\mathcal{E}^{\mathsf{ST}}$ equilibrium with the patronizing equilibrium in Proposition 4. Under the IAA equilibrium, all white workers, including the poor, have a lower utility, though rich white workers acquire a greater skill. Further, all black workers not only have a higher utility, but also acquire more skill.

(c) Compare the $\mathcal{E}^{\mathsf{IM}}$ equilibrium with the patronizing equilibrium in Proposition 4. Under the IAA equilibrium, all white workers, including the poor, have a lower utility, though rich white workers acquire higher skills. Further, all black workers have a higher utility, but while rich black workers acquire less skill, poor black workers acquire greater skills.

6 Comparing class based and identity based affirmative action

We then compare the utilities of the workers under the color-blind CAA equilibrium with the patronising equilibrium under IAA. This would also throw some light on which groups are likely to support a move from IAA to CAA, and which ones are likely to oppose such a move.

We begin by showing that, relative to the patronizing IAA equilibrium, all rich black workers are worse off under the colour-blind CAA equilibrium, whereas all poor white workers are better off. This is intuitive as a switch from IAA to CAA takes away the 'favored' status of the rich black workers, whereas it confers a favored status on the poor whites.

Proposition 6. *Let assumptions 1, 2, 3 and 4 hold. Consider the colour-blind CAA equilibrium in Proposition 2 and the patronizing IAA equilibrium in Proposition 4.*

(a) All rich black workers prefer affirmative action to be identity, rather than class based.

(b) All poor white workers prefer affirmative action to be class, rather than identity based.

Proof. Let $< s^*_{ij}, \pi^*_{ij}, \rho^*_{ij} >$ denote an equilibrium under identity based affirmative action, while $< s^{**}_{ij}, \pi^{**}_{ij}, \rho^{**}_{ij} >$ denotes an equilibrium under class-based affirmative action, where $i = B, W$, and $j = R, P$.

Recall that s' (respectively s'') is the minimum s such that $EE(0)$ intersects WW_R (respectively WW_P). Next note that $s^{**}_{BR} > s' > s^*_{BR}$ since $EE_B(\gamma)$ shifts down under identity-based affirmative action, and $EE_R(\gamma)$ shifts up under class based affirmative action, and $s^{**}_{WP} < s'' < s^*_{WP}$ since $EE_W(\gamma)$ shifts up under IAA, while $EE_P(\gamma)$ shifts down under CAA. The result now follows from Proposition 1. □

Comparing the outcomes for the poor black and the rich white workers is more interesting. This is because poor black workers of course have 'favored' status under both IAA and CAA, whereas the rich white workers are not favored under either. Thus it is not obvious as to how a move from one form of affirmative action to another would affect these two groups.

We next introduce a regularity condition that we call *group MLRP* that extends the notion of MLRP to the group context.

Assumption 5. *Group MLRP: $\frac{1-G(\beta(s))}{G(\beta(s))}\phi(s)$ and $\frac{1-G(\hat{\beta}(s))}{G(\hat{\beta}(s))}\phi(s)$ are both decreasing in s for all s.*

We recall that MLRP ensures that, at the individual level, a higher signal suggests that an agent is more likely to be qualified. At the group level however the informativeness of the signal also depends on the fraction of workers getting skilled. To capture this aspect note that

$$\frac{1-G(\beta(s))}{G(\beta(s))}\phi(s) = \frac{\text{[The fraction of workers remaining unskilled]} \times f_u(s)}{\text{[The fraction of workers getting skilled]} \times f_q(s)}.$$

Thus whenever $\frac{1-G(\beta(s))}{G(\beta(s))}\phi(s)$ is decreasing in s, this ensures that whenever rich workers take their skilling decisions optimally, a higher signal continues to be informative.

How restrictive is assumption 5? One can find simple examples such that assumption 5 holds for all s; it holds, for example, if (a) $f_u(s) = 2$ for all $s \in [0, 1/2]$ and $f_u(s) = 0$, $\forall s \in (1/2, 1]$, and (b) $f_q(s) = 1$ for all $s \in [0, 1]$.[22] More generally, note that $\frac{1-G(\beta(s))}{G(\beta(s))}\phi(s)$ and $\frac{1-G(\hat{\beta}(s))}{G(\hat{\beta}(s))}\phi(s)$ are both necessarily decreasing for all $s < \tilde{s}$. Thus, even in the absence of this assumption, Proposition 7 below goes through whenever the equilibria under consideration all involve cutoff signals that are less than $\tilde{s}$.

We next show that a move from IAA to CAA helps the poor black workers, though it hurts the rich whites.

Proposition 7. *Let assumptions 1, 2, 3, 4, and 5 hold. Consider the colour-blind CAA equilibrium in Proposition 2 and the patronizing IAA equilibrium in Proposition 4.*

[22]For $\frac{1-G(\beta(s))}{G(\beta(s))}\phi(s)$ to be decreasing for all s, it is necessary and sufficient that $\frac{G(\beta(s))(1-G(\beta(s)))(f'_u(s)f_q(s)-f'_q(s)f_u(s))}{wg(.)} < f_u(s)f_q(s)(f_u(s) - f_q(s))$. This condition is sufficient to ensure that $\frac{1-G(\hat{\beta}(s))}{G(\hat{\beta}(s))}\phi(s)$ is decreasing for all s.

(a) *All poor black workers, who are favored under both forms of affirmative action prefer affirmative action to be class, rather than identity based.*

(b) *All rich white workers, who are not part of the favored group under either form of affirmative action, prefer identity based to class-based affirmative action.*

Proof. Let $< s^*_{ij}, \pi^*_{ij}, \rho^*_{ij} >$ denote the patronising equilibrium under identity based affirmative action, while $< s^{**}_{ij}, \pi^{**}_{ij}, \rho^{**}_{ij} >$ denotes the color blind equilibrium under class-based affirmative action, where $i = B, W$, and $j = R, P$. We can mimic the argument in Proposition 6 to show that $s^{**}_{BR} > s^*_{BR}$ and $s^{**}_{WP} < s^*_{WP}$. Hence, from Assumption 3, we have that $\rho^{**}_{BR} < \rho^*_{BR}$ and $\rho^{**}_{WP} > \rho^*_{WP}$.[23]

We next argue that $\rho^{**}_{BP} > \rho^*_{BP}$. Suppose to the contrary that $\rho^{**}_{BP} \leq \rho^*_{BP}$. This implies that $s^{**}_{BP} \geq s^*_{BP}$ (from assumption 3). Given (16), (26) and assumption 5, i.e. the fact that $\frac{1-G(\hat{\beta}(s))}{G(\hat{\beta}(s))}\phi(s)$ is decreasing in s, this implies that $\gamma^* \geq \gamma^{**}$. This in turn implies that $s^{**}_{WR} \leq s^*_{WR}$ (from (19), (29) and assumption 5), and hence $\rho^{**}_{WR} \geq \rho^*_{WR}$ (from assumption 3). We then note that

$$\begin{aligned} \mu_{WR}\rho^*_{WR} + \mu_{WP}\rho^*_{WP} &< \mu_{WR}\rho^{**}_{WR} + \mu_{WP}\rho^{**}_{WP} \\ &= [\mu_{WR} + \mu_{WP}]\rho^{**}_{WP} \text{ (since } \rho^{**}_{WR} = \rho^{**}_{WP}) \\ &= [\mu_{BR} + \mu_{BP}]\rho^{**}_{WP} \text{ (since } \mu_{BP} + \mu_{BR} = 1 = \mu_{WP} + \mu_{WR}) \\ &= \mu_{BR}\rho^{**}_{BR} + \mu_{BP}\rho^{**}_{BP} \text{ (since } \rho^{**}_{WP} = \rho^{**}_{BR} = \rho^{**}_{BP}) \\ &< \mu_{BR}\rho^*_{BR} + \mu_{BP}\rho^*_{BP}, \end{aligned} \tag{34}$$

where the first inequality follows since $\rho^{**}_{WR} \geq \rho^*_{WR}$ and $\rho^{**}_{WP} > \rho^*_{WP}$, and the final inequality follows since $\rho^*_{BP} \geq \rho^{**}_{BP}$ and $\rho^*_{BR} > \rho^{**}_{BR}$. Note however that (34) contradicts the affirmative action constraint under identity-based affirmative action. Hence, $\rho^{**}_{BP} > \rho^*_{BP}$, and consequently from assumption 3, $s^{**}_{BP} < s^*_{BP}$, and hence we have that $\gamma^* < \gamma^{**}$.

Next, given that $\gamma^* < \gamma^{**}$, from (19), (29) and assumption 5, we have that $s^{**}_{WR} > s^*_{WR}$, whereas from (16), (26) and assumption 5, we have that $s^{**}_{BP} < s^*_{BP}$. The result now follows from Proposition 1. □

The intuition comes from the fact that acquiring education is costlier for the poor, so that ceteris paribus they are less likely to get skilled relative to the rich. Consequently, given their income heterogeneity, on the average black workers are more likely to acquire an education relative to the poor. This ensures that the shadow price of equality is higher under a class-based affirmative action compared to that under identity based affirmative action. Consequently, poor blacks, who gain from both forms of affirmative action prefer it to be class-based, since they benefit from the higher shadow price under a CAA. For the same reason, rich white workers prefer IAA.

[23]Note that the comparison is across the same income class, so that we can, for example, invoke the fact that $\hat{\rho}(s)$ is negatively sloped while arguing that, the fact that $s^{**}_{BR} > s^*_{BR}$, implies that $\rho^{**}_{BR} < \rho^*_{BR}$.

6.1 An economy with no rich black workers and 'large' number of poor black and poor white workers

This sub-section is motivated by the following question. Consider an economy that is in the $\mathcal{E}^P$ equilibrium to begin with, so that the poor, irrespective of colour, acquire no skill and are never assigned to task 1. It is arguable that in such a scenario the salient issue is to find a policy where all poor workers, irrespective of identity, benefit. Does there exist economies such that CAA 'ensures' that all the poor are better off, while IAA may fail to do so?

To that end we consider an economy where inequality among blacks is severe, in fact assume that there are no rich black workers i.e. $\lambda_{BR} = 0$. Moreover, the poor include a mix of both black and white workers, neither being too large or too small, formally $\max\{\frac{\lambda_{BP}}{\lambda_P}, \frac{\lambda_{WP}}{\lambda_P}\} < \hat{\rho}(s''')$, where s''' is the maximal s such that EE and WW_R intersect.

We first show that under CAA, in any non-trivial equilibrium all poor workers have a positive probability of being assigned to task 1. Moreover, relative to $\mathcal{E}^P$, all poor workers acquire a greater level of skill and have a higher utility.

Proposition 8. *Suppose assumptions 1, 2, 3, and 4 hold. Moreover, assume that* $\lambda_{BR} = 0$, *and* $\max\{\frac{\lambda_{BP}}{\lambda_P}, \frac{\lambda_{WP}}{\lambda_P}\} < \hat{\rho}(s''')$.

(A) Then in any non-trivial CAA equilibrium with $\gamma^{**} > 0$, *we have that* $\rho^{**}_{BP}, \rho^{**}_{WP} > 0$.

(B) Under any non-trivial CAA equilibrium with $\gamma^{**} > 0$, *the poor have a higher utility and acquire more skills relative to* $\mathcal{E}^P$.

Proof. (A) To begin with, note that $\lambda_{BR} = 0$, so that the CAA constraint entails

$$\frac{\lambda_{BP}}{\lambda_P}\rho^{**}_{BP} + \frac{\lambda_{WP}}{\lambda_P}\rho^{**}_{WP} = \rho^{**}_{WR}.$$

Suppose to the contrary that $\rho^{**}_{BP} = 0$, say. Then

$$\frac{\lambda_{BP}}{\lambda_P}\rho^{**}_{BP} + \frac{\lambda_{WP}}{\lambda_P}\rho^{**}_{WP} = \frac{\lambda_{WP}}{\lambda_P}\rho^{**}_{WP} < \hat{\rho}(s''') \le \rho^{**}_{WR}, \tag{35}$$

where the first equality follows since $\rho^{**}_{BP} = 0$, the first inequality follows from our proposition hypothesis that $\max\{\frac{\lambda_{WP}}{\lambda_P}, \frac{\lambda_{BP}}{\lambda_P}\} < \hat{\rho}(s''') \le \rho^{**}_{WR}$, and the final inequality follows since $\gamma^{**} > 0$ and $\rho^{**}_{WR} \neq 0$ (otherwise we have a trivial equilibrium). A similar argument applies if $\rho^{**}_{WP} = 0$.

(B) The result follows since from step (a) we have that $s^{**}_{BP}, s^{**}_{WP} < 1$, and the fact that under $\mathcal{E}^P$, we have that $\bar{s}_{BP} = \bar{s}_{WP} = 1$. □

Note that we restrict attention to equilibria that has the natural property that the shadow price of the affirmative action constraint i.e. γ^{**} is positive. One sufficient condition that ensures that this is indeed the case is given by

Assumption 6.

(a) $wg'(\beta(s))[f_u(s) - f_q(s)]^2 + g(\beta(s))[f'_u(s) - f'_q(s)] < 0$ *and* $wg'(\hat{\beta}(s))[f_u(s) - f_q(s)]^2 + mg(\hat{\beta}(s))[f'_u(s) - f'_q(s)] < 0$.

(b) The likelihood ratio function $\phi(s)$ *is sufficiently convex, i.e.* $\phi(s)\phi''(s) > 2(\phi'(s))^2$.[24]

Given the same parameter configurations we then argue that under identity-based affirmative action and severe poverty i.e. large m , poor white workers are never assigned to task 1.

Proposition 9. *Suppose assumptions 1, 2, 3, and 4 hold. Moreover, let* $\lambda_{BR} = 0$*. If* $EE_W(\lambda_B(x_u + w))$ *does not intersect* WW_P*, and* m *is large, then in* any *CAA equilibrium with* $\gamma^* > 0$ *the poor white are never assigned to task 1.*

Proof. From (9), fixing s_{iP}, we have that

$$\lim_{m\to\infty} \pi_{iP} = \lim_{m\to\infty} G(\frac{\beta(s_{iP})}{m}) = G(0) = 0.$$

Thus $\lim_{m\to\infty} \gamma^*(m) = \lambda_B(x_u + w)$, since otherwise for m sufficiently large $EE_B(\gamma)$ and WW_P will not intersect (from (26) and (27)). Consequently, for m sufficiently large $EE_W(\gamma)$ and WW_P will not intersect, so that $\rho^*_{WP} = 0$. □

7 Conclusion

We examine both class, as well as identity based affirmative action in a framework that allows for heterogeneity in identity, as well as income, and where, moreover, poverty is more of a concern for black workers. We establish that with class-based affirmative action there exists an equilibrium that is colour-blind, i.e. the poor face the same cut-offs irrespective of their identity, as do the rich. Whereas, if affirmative action is identity-based, then there exists an equilibrium that has elements of patronization in that black workers of both income classes face lower cutoffs relative to their white counterparts. We find that relative to the patronizing equilibrium under IAA, poor black workers prefer the color-blind CAA equilibrium though they benefit from affirmative action under both IAA and CAA, whereas the rich white workers prefer IAA though they are not protected under either form of affirmative action. In fact, the relative preference for the two policies are completely class-based, with all poor workers preferring CAA, whereas all rich workers prefer IAA. Moreover, we find that there exist economies where CAA ensure that all poor workers are assigned to the skilled task with positive probability, whereas under IAA that is not necessarily true.

Given this diversity in preferences, the implementation of either IAA or CAA would therefore depend on the economic and political clout enjoyed by the various groups. Not surprisingly the choice of IAA versus CAA is not just very complex, but also an extremely divisive political issue.

[24] The argument involves showing that under Assumption 6, WW_P and WW_R are concave and $EE_k(\gamma)$, $k \in \{P, R, B, W\}$, is convex in s. This assumption holds, for example, for the following set of functions: (a) c is uniformly distributed over $[0, 1]$, so that $g(.) = 1$, $g'(s) = 0$, (b) $\phi(s) = e^{-2s^{\frac{1}{3}}} = f_u(s)$ for all $s \in [0, 1]$, and (c) $f_q(s) = 1$ for all $s \in [0, 1]$. The proof is available with the authors.

8 Appendix

Proof of Proposition 3. (a) From the definition of $\mathcal{E}^{P}$, recall that

$$\bar{s}_{BP}(\mathcal{E}^{P}) = \bar{s}_{BR}(\mathcal{E}^{P}) = \bar{s}_{WP}(\mathcal{E}^{P}) = 1 > \bar{s}_{WR}(\mathcal{E}^{P}).$$

Whereas from Proposition 2, under the colour-blind CAA equilibrium one has that

$$1 > s^{**}_{BR} = s^{**}_{WR} > s^{**}_{BP} = s^{**}_{WP}.$$

Note that $\bar{s}_{BP}(\mathcal{E}^{P}) = \bar{s}_{BR}(\mathcal{E}^{P}) = \bar{s}_{WP}(\mathcal{E}^{P}) = 1 > s^{**}_{BR} > s^{**}_{BP} = s^{**}_{WP}$. Further, $s^{**}_{WR} > \bar{s}_{WR}(\mathcal{E}^{P})$ (since $\gamma^{**} > 0$). Thus from Proposition 1, rich white workers are worse off, while all other workers are better off.

Skill acquisition by all poor workers, as well as rich black workers increase under CAA since they acquire no skill under $\mathcal{E}^{P}$. Skill acquisition by the rich white workers increases as well since EE_R intersects WW_R to the left of $\tilde{s}$. Consequently, the firm's profit increases as well.

(b) From the definition of $\mathcal{E}^{ST}$, recall that

$$\bar{s}_{BP}(\mathcal{E}^{ST}) = \bar{s}_{BR}(\mathcal{E}^{ST}) = 1 > \bar{s}_{WP}(\mathcal{E}^{ST}) > \bar{s}_{WR}(\mathcal{E}^{ST}).$$

Whereas from Proposition 2, under the colour-blind CAA equilibrium one has that

$$1 > s^{**}_{BR} = s^{**}_{WR} > s^{**}_{BP} = s^{**}_{WP}.$$

Note that $\bar{s}_{BP}(\mathcal{E}^{ST}) = \bar{s}_{BR}(\mathcal{E}^{ST}) = 1 > s^{**}_{BR} > s^{**}_{BP}$ so black workers, both rich and poor are subject to a lower standard. Thus from Proposition 1, they are better off under CAA. As to white workers, $s^{**}_{WP} < \bar{s}_{WP}(\mathcal{E}^{ST})$ and $s^{**}_{WR} > \bar{s}_{WR}(\mathcal{E}^{ST})$ (since $\gamma^{**} > 0$). Thus from Proposition 1, the rich (respectively the poor) white workers are worse off (respectively better off) under CAA.

That skill acquisition by black workers increase under CAA is clear since they acquire no skill under $\mathcal{E}^{ST}$. Whereas skill acquisition by the poor white workers decrease since $\gamma^{**} > 0$ and WW_P is increasing for $s < \tilde{s}$. By a similar argument skill acquisition by rich white workers increases.

(c) From the definition of $\mathcal{E}^{IM}$, recall that

$$\bar{s}_{BP}(\mathcal{E}^{IM}) = 1 > \bar{s}_{WP}(\mathcal{E}^{IM}) > \bar{s}_{WR}(\mathcal{E}^{IM}) = \bar{s}_{BR}(\mathcal{E}^{IM}).$$

Whereas from Proposition 2, under the colour-blind CAA equilibrium one has that

$$1 > s^{**}_{BR} = s^{**}_{WR} > s^{**}_{BP} = s^{**}_{WP}.$$

Note that $\bar{s}_{BP}(\mathcal{E}^{IM}) = 1 > s^{**}_{BP}$ so that poor black workers are better off under CAA. Further, $s^{**}_{WP} < \bar{s}_{WP}(\mathcal{E}^{IM})$, $s^{**}_{WR} = s^{**}_{BR} > \bar{s}_{WR}(\mathcal{E}^{IM}) = \bar{s}_{BR}(\mathcal{E}^{IM})$ (since $\gamma^{**} > 0$). Thus from Proposition 1, rich workers, both black and white, are worse off, while poor workers, both black and white, are better off.

That skill acquisition by poor black workers increase under CAA is clear since they acquire no skill under $\mathcal{E}^{ST}$. Whereas skill acquisition by the poor white workers decreases, and that by rich white workers increases since $\gamma^{**} > 0$ and WW_P and WW_R are increasing for $s < \tilde{s}$. $\square$

Proof of Proposition 4. (a) Define $\hat{s}_{ij}$ as in Proposition 2, and arguing as in Proposition 2, we find that that $\hat{s}_{BR}(0) = \hat{s}_{WR}(0) = s' < \hat{s}_{BP}(0) = \hat{s}_{WP}(0) = s'' < \tilde{s}$. Further, equation (22) holds, so that $\hat{\rho}(s') > \hat{\hat{\rho}}(s'')$.

To show that there exists $\gamma^*(m) > 0$, such that $\hat{s}_{ij}(\gamma^*)$ satisfies the identity-based affirmative action constraint (3). As in Proposition 2, we shall invoke the intermediate value theorem to establish our claim. To that end we define

$$D(\gamma) = \mu_{BR}\hat{\rho}(\hat{s}_{BR}(\gamma)) + \mu_{BP}\hat{\hat{\rho}}(\hat{s}_{BP}(\gamma)) - \mu_{WR}\hat{\rho}(\hat{s}_{WR}(\gamma)) - \mu_{WP}\hat{\hat{\rho}}(\hat{s}_{WP}(\gamma)),$$

and let γ increase from zero. Note that

$$\begin{aligned} D(0) &= \mu_{BR}\hat{\rho}(\hat{s}_{BR}(0)) + \mu_{BP}\hat{\hat{\rho}}(\hat{s}_{BP}(0)) - \mu_{WR}\hat{\rho}(\hat{s}_{WR}(0)) - \mu_{WP}\hat{\hat{\rho}}(\hat{s}_{WP}(0)) \\ &= \mu_{BR}\hat{\rho}(\hat{s}_{BR}(0)) + (1-\mu_{BR})\hat{\hat{\rho}}(\hat{s}_{BP}(0)) - \mu_{WR}\hat{\rho}(\hat{s}_{BR}(0)) - (1-\mu_{WR})\hat{\hat{\rho}}(\hat{s}_{BP}(0)) \\ &= \mu_{WR}(\hat{\hat{\rho}}(\hat{s}_{BP}(0)) - \hat{\rho}(\hat{s}_{BR}(0))) - \mu_{BR}(\hat{\hat{\rho}}(\hat{s}_{BP}(0)) - \hat{\rho}(\hat{s}_{BR}(0))) \\ &= [\mu_{WR} - \mu_{BR}][\hat{\hat{\rho}}(s'') - \hat{\rho}(s')] < 0, \end{aligned}$$

where the last inequality follows from (22), and the fact that from Assumption 4, $\mu_{WR} = 1 - \mu_{WP} > 1 - \mu_{BP} = \mu_{BR}$. We then observe that $\hat{s}_{ij}(\gamma)$ is continuous, where the continuity of $\hat{s}_{Wj}(\gamma)$ follows since from assumption 2(b), $EE_W(\gamma)$ and WW_R always have an intersection for all $\gamma \leq \lambda_B(x_u + w)$ (we mimic the argument in footnote (18) to establish this claim), with $\hat{s}_{BR}(\gamma)$, $\hat{s}_{BP}(\gamma)$ being decreasing, and $\hat{s}_{WR}(\gamma)$, $\hat{s}_{WP}(\gamma)$ being increasing in γ.

Further, as $\gamma \to \lambda_B(x_u + w)$, $\hat{s}_{Bj}(\gamma)$ goes to zero (from (26) and (27)), so that $\hat{\rho}(s_{BR})$ and $\hat{\hat{\rho}}(s_{BP})$ go to 1, and consequently $[\mu_{BR}\hat{\rho}(\hat{s}_{BR}(\gamma)) + \mu_{BP}\hat{\hat{\rho}}(\hat{s}_{BP}(\gamma))]|_{\gamma\to\lambda_B(x_u+w)} = 1$. Given that $\hat{\rho}(\hat{s}_{WR}(\lambda_B(x_u + w))) < 1$, and $\hat{\hat{\rho}}(\hat{s}_{WP}(\lambda_B(x_u + w))) < 1$ (since $\hat{s}_{Wj}(\lambda_B(x_u + w)) > 0$), we have that $D(\gamma)|_{\gamma\to\lambda_B(x_u+w)} > 0$. Consequently, given that $D(0) < 0$, from the continuity of $D(\gamma)$ there exists $\gamma^* > 0$ such that $D(\gamma^*) = 0$. Finally, note that by construction, $\hat{s}_{Bj} < \hat{s}_{Wj}$, $j \in \{P, R\}$, so that the equilibrium is non-stereo-typing.

Local stability again follows from the construction of this equilibrium.

(b) Suppose to the contrary a colour-blind equilibrium exists where $\rho^*_{BP} = \rho^*_{WP}$ and $\rho^*_{BR} = \rho^*_{WR}$. Then the IAA constraint simplifies to

$$\mu_{BP}\rho^*_{BP} + \mu_{BR}\rho^*_{BR} = \mu_{WP}\rho^*_{BP} + \mu_{WR}\rho^*_{BR}.$$

Given that $\rho^*_{BP} < \rho^*_{BR}$, and $\mu_{BP}+\mu_{BR} = 1 = \mu_{WP}+\mu_{WR}$, this cannot hold unless $\mu_{BP} = \mu_{WP}$. $\square$

Proof of Proposition 5. (a) From the definition of $\mathcal{E}^P$, recall that

$$\bar{s}_{BP}(\mathcal{E}^P) = \bar{s}_{BR}(\mathcal{E}^P) = \bar{s}_{WP}(\mathcal{E}^P) = 1 > \bar{s}_{WR}(\mathcal{E}^P).$$

Whereas from Proposition 4, under the patronizing IAA equilibrium one has that

$$s^*_{BR} < s^*_{WR}, \ \ s^*_{BP} < s^*_{WP}.$$

Note that $s^*_{BR}, s^*_{BP} < 1 = \bar{s}_{BP}(\mathcal{E}^P) = \bar{s}_{BR}(\mathcal{E}^P)$, so that all black workers are better off. Further, since $\gamma^* > 0$, $s^*_{WR} > \bar{s}_{WR}(\mathcal{E}^P)$, so that rich white workers are worse off, though they acquire more skill since the minimum s such that EE_W intersects WW_R is less than $\tilde{s}$. Whereas $s^*_{WP} \leq \bar{s}_{WP}(\mathcal{E}^P)$, so that poor white workers are not worse off.

Finally, given that black workers acquire no skill under $\mathcal{E}^P$, skill acquisition increases under IAA.

(b) From the definition of $\mathcal{E}^{ST}$, recall that

$$\bar{s}_{BP} = \bar{s}_{BR} = 1 > \bar{s}_{WP} > \bar{s}_{WR}.$$

Whereas from Proposition 4, under the patronizing IAA equilibrium one has that

$$s^*_{BR} < s^*_{WR}, \ \ s^*_{BP} < s^*_{WP}.$$

Note that $\bar{s}_{BP} = \bar{s}_{BR} = 1 > s^*_{BP} > s^*_{BR}$ so black workers, both rich and poor are subject to a lower standard. Thus from Proposition 1, they are better off under IAA. As to white workers, $s^*_{WP} > \bar{s}_{WP}$ and $s^*_{WR} > \bar{s}_{WR}$ (since $\gamma^* > 0$). Thus from Proposition 1, both rich and poor white workers are worse off under IAA. Skill acquisition by rich white workers increases since the minimum s such that EE_W intersects WW_R is less than $\tilde{s}$.

Given that black workers acquire no skill under $\mathcal{E}^{ST}$, but a positive amount under IAA, skill acquisition by black workers increase under IAA.

(c) From the definition of $\mathcal{E}^{IM}$, recall that

$$\bar{s}_{BP} = 1 > \bar{s}_{WP} > \bar{s}_{WR} = \bar{s}_{BR}.$$

Whereas from Proposition 4, under the patronizing IAA equilibrium one has that

$$s^*_{BR} < s^*_{WR}, \ \ s^*_{BP} < s^*_{WP}.$$

Note that $\bar{s}_{BP} = 1 > s^*_{BP}$ so that poor black workers are better off under IAA. Further, $s^*_{WP} > \bar{s}_{WP}$, $s^*_{WR} > \bar{s}_{WR}$ and $s^*_{BP} < \bar{s}_{BP}$ (since $\gamma^* > 0$). Thus from Proposition 1, white workers, both rich and poor, are worse off, while rich black workers are better off.

That skill acquisition by poor black workers increase under IAA is clear since they acquire no skill under $\mathcal{E}^{IM}$. Whereas skill acquisition by rich black workers decreases since $\gamma^* > 0$ and WW_P is increasing for $s < \tilde{s}$. Skill acquisition by rich white workers increases since the minimum s such that EE_W intersects WW_R is less than $\tilde{s}$. □

9 Bibliography

Abdulkadiroğlu, A. (2005): College Admissions with Affirmative Action. International Journal of Game Theory 33, 535-549.

Abdulkadiroğlu, A. and T. Sönmez (2003): School Choice: A Mechanism Design Approach. American Economic Review 93, 729-747.

Allen, L. (1995): The role model argument and faculty diversity, in S. M. Cahn, ed., The affirmative action debate. New York: Routledge, pp. 121-34.

Alon, S. (2011): The diversity dividends of a need-blind and color-blind affirmative action policy. Social Science Research, 40(6): 1494-1505.

Arrow, K.J (1972): Some models of racial discrimination in the labour market. In Racial Discrimination in Economic Life, Lexington Books.

Banerjee, A., Bertrand, M., Datta, S. and Mullainathan, S. (2009): Labour market discrimination in Delhi: Evidence from a field experiment. Journal of Comparative Economics 37, 14-27.

Becker, G.S. (1957): The Economics of Discrimination. Chicago: University of Chicago Press.

Bertrand, M. and Mullainathan, S. (2004): Are Emily and Greg more emplyable than Lakisha and Jamal? A field experiment on labour market discrimination. American Economic Review 94, 991-101.

Carter, M.R and Lybbert, T.J. (2012): Consumption versus asset smoothing: testing the implications of poverty trap theory in Burkina Faso. Journal of Development Economics 99, 255-264.

Chung, K.S (2000): Role models and arguments for affirmative action. American Economic Review, 90(3): 640-648.

Coate, S. and Loury, G. (1993): Will affirmative action policies eliminate negative stereotypes? American Economic Review, 83(5): 1220-1240.

Deshpande, A. and Ramachandran, R. (2014): How backward are the other backward classes? Changing contours of caste disadvantage in India. Mimeo, Delhi School of Economics.

Fang, H. and Moro, A. (2010): Theories of statistical discrimination and affirmative action: a survey. NBER Working Paper.

Fang, H. and Norman, P. (2006): Government-mandated discriminatory policies: Theory and evidence. International Economic Review 47, 361-389.

Fryer, R.G. (2007): Belief flipping in a dynamic model of statistical discrimination. Journal of Public Economics 91, 1151-1166.

Fryer, R.G. and Loury, G.C. (2013): Valuing Diversity. Journal of Political Economy 121, 747-774.

Hafalir, I.E., Yenmez, M.B., and Yildirim, M.A. (2013): Effective Affirmative Action in School Choice. Theoretical Economics 8, 325-263.

Hnatkovska, V., Lahiri, A. and Paul, S. (2012): Castes and Labor Mobility. American Economic Journal: Applied Economics, American Economic Association 4(2), 274-307.

Holzer, H. and D. Neumark (2000): Assessing affirmative action. Journal of Economic Literature 38, 483-568.

Jackson, B.O (1987): The effects of racial group consciousness on political mobilization in American cities. Western Political Quarterly, 40(4): 631-646.

Kamada, Y. and F. Kojima (2011): Improving Efficiency in Matching Markets with Regional Caps: The Case of the Japan Residency Matching Program. Unpublished manuscript.

Kojima, F. (2012): School Choice: Impossibilities for Affirmative Action. Games and Economic Behaviour 75, 685-693.

Leonard, J.S. (1984): The impact of affirmative action on employment. Journal of Labour Economics 2, 439-464.

Loury, G. (1977): A dynamic theory of racial income distribution. In. P. Wallace and A. LaMond eds. Women, Minorities, and Employment Discrimination. Lexington, M.A.: D.C. Heath, pp. 153-188.

Loury, G. (1981): Is equal opportunity enough? American Economic Review 71, 122-126.

Lundberg, J.S. and R. Startz (1983): Private discrimination and social intervention in competitive labour markets. American Economic Review 73, 340-347.

Lundberg, S.J. (1991): The enforcement of equal opportunity laws under imperfect information: Affirmative action and alternatives. Quarterly Journal of Economics 106, 309-206.

Milgrom, P. and S. Oster (1987): Job discrimination, market forces and the invisibility hypothesis. Quarterly Journal of Economics 102, 453-476.

Moro, A. and Norman, P. (2003): Affirmative action in a competitive economy. Journal of Public Economics 87, 567-594.

Moro, A. and Norman, P. (2004): A general equilibrium model of statistical discrimination. Journal of Economic Theory 114, 1-30.

Nelson, W. and Meranto, P.J (1977): Electing black mayors: political action in the black community. Columbus: Ohio State University Press.

Pathak, P. A. (2011): The Mechanism Design Approach to Student Assignment. Annual Review of Economics 3, 513-536.

Phelps, E.S. (1972): The statistical theory of racism and sexism. American Economic Review 62, 659-61.

Reece, J. (2011): Revisiting class-based affirmative action in government contracting. Washington University Law Review 88, 1309-1352.

Sander, R.H (1997): Experimenting with class-based affirmative action. Journal of Legal Education, 47(4): 472-503.

Smith, J.P. and Welch, F. (1984): Affirmative action and labour markets. Journal of Labour Economics 2, 269-301.

Welch, F. (1976): Employment quotas for minorities. Journal of Political Economy 84, 8105-39.

Welch, F. (1989): Affirmative action and discrimination. In S. Shulman and W. Darrity, Jr., eds., The Question of Discrimination. Middletown, CT: Wesleyan University Press, pp. 153-189.

Westkamp, A. (2013): An Analysis of the German University Admissions System. Economic Theory 53, 561-589.

Wilson, W.J (1978): The declining significance of race. Chicago: University of Chicago Press.