Affirmative Action in the Presence of a Creamy Layer: Identity or Class Based?

Guha, Brishti and Roy Chowdhury, Prabal

Jawaharlal Nehru University, Indian Statistical Institute, New Delhi

2017

Affirmative Action in the Presence of a Creamy Layer: Identity or Class based?

Brishti Guha* Prabal Roy Chowdhury†

Abstract

We analyze affirmative action in the presence of income heterogeneity, with the target group containing a creamy layer, as well as poorer members. We find that a move from identity to class-based affirmative action would affect the creamy layer and the poorer members of the target group differently. While it would help the poorer members of the target group, it would hurt the creamy layer, as well as the non-target group. Further, temporary affirmative action need not have a permanent effect, since removing identity-based affirmative action may harm the poorer members of the target group. Thus either removing identity-based affirmative action, or switching to a class-based affirmative action is likely to be politically difficult.

JEL No.: J71, J78, D78, D82.
Key Words: Affirmative action; creamy layer; class-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

Many countries adopt affirmative action policies in favour of traditionally underprivileged 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.[1] In this paper we focus on an aspect of affirmative action that has been relatively under-researched in the literature, namely the fact that the intended beneficiaries typically contain people with very different income levels, including the so called "creamy layer". Furthermore, over time more and more members of the target groups have been entering the creamy layer, so that the issue of income heterogeneity among the target group is gaining in relevance.[2]

We compare and contrast two alternative forms of affirmative action, identity and class-based, in the presence of income heterogeneity. While affirmative action has traditionally been identity-based, class-based affirmative action is being debated, as well as implemented (at least in certain sectors), in several countries. 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' (henceforth OBC) beneficiary with income exceeding the cutoff is excluded.[3] 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 socioeconomically disadvantaged areas, whose employees

[1] 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] Poverty levels among African-Americans in the US, for example, have fallen sharply over the last 50 years, from 41.8% in 1966, to 27.2% in 2012 (Pew Research Center, January 2014). In India, poverty rates among the scheduled castes, beneficiaries of affirmative action, fell from 62.4% in 1993-94, to 31.5% in 2011-12. Similarly, poverty rates among the other beneficiaries of Indian affirmative action, namely scheduled tribes and other backward castes (OBCs), declined by over 17 percentage points between 2003-04 and 2011-12, as opposed to an 11.6% decline for non-target groups (Business Standard, March 14, 2014, and also Times of India, May 5, 2015). Hnatkovska et al. (2012) and Deshpande and Ramachandran (2014) examine divergence along other dimensions as well, and reach conclusions with a similar flavour.

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

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,[4] France,[5] Israel,[6] New Zealand[7] and South Africa.[8]

We are interested in several key questions that are germane to the political economy of affirmative action. In order to pose these issues sharply, we consider a society that is divided along both identity and class lines, in particular we examine a society consisting of three groups of workers, the (rich) whites, the poor blacks, and the creamy layer blacks.[9]

First, we examine the effects of a switch from identity to income-based affirmative action. If the objective of the government is to help the poor blacks, then such a policy may be more effective in achieving this target. While speaking about inequality in India, for example, Thomas Piketty said "... the long-term objective should be to gradually move away from a caste-based reservation system to a system of reservation that is more based on parental income, parental wealth."[10] Further, given that income disparity among the target groups are declining in several countries, some of the gains from affirmative action may go to the increasingly affluent and increasingly larger creamy layer blacks as long as it is identity based. How would such a switch to income based affirmative action affect the utility of the various groups?[11] Answering

[4]In Brazil, quotas for the poor coexist 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).

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

[6]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).

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

[8]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).

[9]This is not to say that there is no income heterogeneity among 'whites', but just that poverty is relatively more of a concern among the 'blacks'.

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

[11]Interestingly, sometimes class has been used as proxy for identity based affirmative action. In the US, when colleges were forbidden from explicit use of race-based quotas for admissions, some

this question would help us understand how different groups are going to react to any such attempted switch.

Second, following Coate and Loury (1993), one can ask if a temporary affirmative action policy is likely to have a permanent effect or not. This is important given that continuation of affirmative action over a long period can have serious costs, including a possible backlash from the non-target groups.[12] In particular is it possible that unlike in Coate and Loury (1993) removal of affirmative action can generate a worse labour market outcome for the poor black workers? If so, then once enforced, affirmative action policies may be difficult to remove as doing so would entail serious political costs.

Following Arrow (1972) and Coate and Loury (1993), we examine a model where the employers receive an imperfect signal regarding the skill level of all workers. Further, the poor black workers have higher opportunity costs of investing in education relative to the whites, as well as the creamy layer blacks. This may arise because of various reasons that we discuss later in the paper. The belief of the firms regarding any worker is conditioned both by this signal, as well as their perception regarding the average skill level of the concerned group. Based on this belief they assign workers to either a skilled job (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. Consequently, the equilibrium may display statistical discrimination and stereo-typing.

Next turning to the results, we demonstrate that there can be multiple equilibria, and examine conditions under which an equilibrium with stereo-typing, and even patronization (whereby the target groups are held to a lower standard) may, or may not exist. We also establish that equilibrium utility levels are closely tied to labour market outcomes.

The interaction between identity based stereo-typing and target group heterogeneity gives rise to interesting new polarisations. Critically, the investment decisions, and consequently the interests, of the members of the creamy layer and the white workers, may get aligned to a certain extent, since both these groups would find their investment decisions relatively less affected by income constraints vis-a-vis the poor black workers. This in turn may mean that the interests of the poor blacks

---

colleges began using "class-based" affirmative action (Sander 1997). However, following an adverse US Supreme Court ruling in 2003, class-based affirmative action has not been widely implemented.

[12]In the Indian state of Gujarat, for example, the relatively wealthy and upwardly mobile community of "Patels" have started an agitation for OBC status, which would entitle them to benefit from affirmative action (The Hindu, 26th August, 2015). The agitation not only led to arson and destruction of public property, but also widespread violence and even death. Similarly, the Jat community in Haryana, and the Gujjars in Rajasthan, are also demanding OBC status in India.

differ from those of the creamy layer, even though both benefit from affirmative action. We then argue that any attempt to either repeal such policies, or to replace identity based policies with class based ones, may face political obstacles as these may worsen the labour market outcome for one group or the other.

We find that if an income based affirmative action policy were to replace an identity-based one, then all creamy layer blacks are worse off in that they are assigned to the skilled task with its higher wages at a lower rate. This in turn implies that all creamy layer black workers have a greater utility under income-based affirmative action. This is intuitive given that creamy layer blacks will cease to be beneficiaries if affirmative action is class based. Interestingly, given some regularity conditions, all poor black workers fare strictly better, while all white workers are strictly worse off in terms of both labour market outcomes and their utility.

This suggests that if the objective of affirmative action is to help the poor blacks, then class-based affirmative action is more apt. However, the political economy of such a switch is clearly complex. The analysis suggests that blacks may find it difficult to present a unified stand in support of such a switch, given that creamy layer blacks are unlikely to support such a move. Of course, white workers are also unlikely to favour such a move.[13]

In fact, even if the outcome involves patronisation, so that firms hold poor black workers to a lower standard because of pessimistic beliefs about their skills, we find that creamy layer blacks still fare worse under class-based affirmative action. The interests of whites and poor blacks would diverge in that one group benefits and one group gets hurt by a transition from an identity-based to a class-based policy. Thus, irrespective of whether the equilibrium involves patronisation or not, such a policy switch is unlikely to find support from all groups.

Moreover, unlike Coate and Loury (1993), we find that the removal of affirmative action, whether identity-based or income-based, can generate a worse labour market outcome for poor black workers. This is more likely if poor black workers form a significant proportion of all black workers, suggesting that once enforced, affirmative action policies may be difficult to remove as doing so would entail serious political costs. Thus temporary affirmative action policy is unlikely to have a permanent effect.

Finally, we examine how the *introduction* of affirmative action affects the three

[13]In this context we note that several authors, including Jackson (1987), Nelson and Meranto (1977) and Wilson (1978), find that blacks may fail to present a united front politically, and some of the fault lines may be class-based. For example, Jackson (1987) finds that a black mayoral candidate received overwhelming support from the poorer segments of the black population, but not from affluent blacks.

groups. We find that the interests of the whites and the creamy layer blacks are opposed. In fact, while the whites are worse off as far as being assigned to the skilled task is concerned, the creamy layer blacks are better off. In this context it is interesting that in the US, for example, Feagin and Porter (1995) discuss instances where whites opposed to affirmative action have formed coalitions with like-minded members of the black community (though they do not specify if these whites allied specifically with creamy layer or poor blacks). This suggests that affirmative action policies may require the creamy layer to gain in size, so that they can add their voice in favour of such action.

1.1 Literature Review

The literature closest to ours is the one on statistical discrimination, pioneered by Phelps (1972) and Arrow (1973).[14] The central idea behind statistical discrimination is that individual 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-a-vis the non-target group.

Turning to the theoretical literature,[15] 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. Finally, Lundberg

[14]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 (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).

[15]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.

(1991) examines the impact of affirmative action when regulators do not observe the firms' personnel policies, while 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 income heterogeneity among the target group, and studying the interactions between identity-based stereotyping, and features that generate higher costs of education for poor blacks (such as credit market imperfections). Further, we then use this framework to examine several issues of interest that have been relatively unexplored in the literature, e.g. a switch to income based affirmative action, among others.

The rest of the paper is organized as follows. In the next section we describe the framework analyzed in this paper and study the outcome without any affirmative action. Section 3 examines the outcome when identity-based affirmative action is present, comparing it with the outcome in the absence of any affirmative action. While section 4 considers class-based affirmative action, in section 5 we examine the implications of replacing identity-based with class-based affirmative action. Finally, Section 6 concludes.

2 The Framework

We consider an economy divided along both identity and class lines, comprising three kinds of workers, λ_W white workers, λ_1 creamy layer black and λ_2 poor black workers, and a large number of firms. The number of workers is normalised to 1, so that $\lambda_W + \lambda_1 + \lambda_2 = 1$. Let λ_B denote the proportion of blacks in the population, and λ_R the proportion of the non-poor, so that $\lambda_B \equiv \lambda_1 + \lambda_2$, and $\lambda_R \equiv \lambda_W + \lambda_1$. All agents are risk neutral.

The workers are randomly matched to firms, with all workers finding a match. Following the matching process, the firms assign workers to either of two tasks, 1 or 2, where task 1 requires skill, whereas task 2 does not. In task 1, the payoff of the principal 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.

For all groups, acquiring the requisite skill is costly. This cost is c for the white and creamy layer blacks, and cm (where $m > 1$) for poor blacks, where c is idiosyncratic and distributed over $[0, \infty)$ according to the distribution function $G(c)$ (and density function $g(c)$), where $G(c)$ is continuously differentiable and identical for all three groups. That poor black workers face higher costs could be because of several reasons. It could possibly reflect the fact that a poor black 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 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 costlier for the poor (as it makes consumption smoothing harder).[16] Alternatively, this could arise because with imperfect credit markets, the poor black workers may face higher interest rates in the credit market if they want to fund skill acquisition. In fact there is some evidence that poor students of colour take on more debt, and at higher interest rates, than their counterparts (according to https://www.americanprogress.org/issues/race/news/2013/05/16/63533/borrowers-of-color-need-more-options-to-reduce-their-student-loan-debt/). Finally, it could be because her environment adversely affects a poor black worker's ability to get skilled.[17]

For every worker assigned to firms, the firms can observe group identity, but they cannot observe whether the worker has acquired the necessary skill. The firms however do observe a signal regarding their skill level. 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. Finally, let

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

be well defined for all s, and positive for all $1 > s > 0$. 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* $\lim_{s \to 1} \phi(s) = 0$.

We then specify the utility function of the firms, as well as the workers. The

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

[17] One such cost could arise because the poor blacks may lack role models. As argued by Allen (1995) and Chung (2000), this could be because role models may provide critical information to poor black workers, as well as mentor them, apart from acting as a source of inspiration. For creamy layer black workers this issue can be less important, since with greater income, and consequently greater access and homogenization, they may possibly draw their role model from among the whites as well.

payoff to a 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 the subscript i denote group identity, and noting that $c_i = c$ for whites and creamy layer blacks $(i = W, 1)$, while $c_2 = cm$ for poor blacks, the utility of a group i worker with cost of c_i is

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

The timeline is as follows. Nature moves first, choosing the level of c for every worker. The workers themselves get to observe their own level of c, but the firms do not. Then the workers decide whether to acquire the skill required for task 1, or not. In the next stage the workers are matched to firms, when a firm gets to observe a signal regarding the skill level of a worker assigned to it. Finally, the firms decide on task allocation.

2.1 Equilibrium in the Absence of any Affirmative Action

We first analyze the benchmark case in the absence of affirmative action. Note that in this case the outcomes for the three groups can be examined separately.

The firms' decisions: Consider a firm facing a worker of group i who emits a signal s. If the firm believes that a proportion π_i 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

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

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. $B(\pi_i, s)(x_q - w) - (1 - B(\pi_i, s))(x_u + w) \geq 0$, i.e.

$$r \equiv \frac{x_q - w}{x_u + w} \geq \frac{1 - \pi_i}{\pi_i} \phi(s). \tag{2}$$

Given assumption 1, the firms' decision is characterised by a cutoff s_i such that all workers with a signal greater than s_i are assigned to task 1, where, for all $\pi_i < 1$, s_i is well defined and solves:

$$r \equiv \frac{x_q - w}{x_u + w} = \frac{1 - \pi_i}{\pi_i} \phi(s_i). \tag{3}$$

From MLRP it follows that s_i is decreasing in π. Consequently the graph of $s_i(\pi)$, call it EE, is negatively sloped in $s - \pi$ space.

Let $\rho_i(s_i, \pi_i)$ denote the probability that a randomly drawn worker from group i is assigned to task 1, given that the cutoff signal for this group is s_i, i.e.

$$\rho_i(s_i, \pi_i) = \pi_i[1 - F_q(s_i)] + (1 - \pi_i)[1 - F_u(s_i)]. \tag{4}$$

The workers' decision: Next consider the decision problem facing a worker of type i, who believes that firms will assign her to task 1 if and only if she emits a signal of s_i, or higher. Note that acquiring the skill increases a worker's chances of getting assigned to task 1, as she is more likely to send a signal greater than s_i. Letting $\beta(s)$ denote the increase in expected gross income from skill acquisition, we can write $\beta(s) \equiv w(F_u(s) - F_q(s))$. It is straightforward to check that $\beta(0) = \beta(1) = 0$, so that $G(\beta(0)) = G(\beta(1)) = 0$. Further, given MLRP, $G(\beta(s)) \geq 0$ and increasing if and only if $\phi(s) > 1$. This in turn implies that $G(\beta(s))$ is single-peaked, attaining a maximum value at some $\tilde{s} > 0$, so that $G(\beta(s))$ is increasing if and only if $s < \tilde{s}$.

First consider white and CLB workers. Such a worker with cost c acquires the skill if and only if $w(1 - F_q(s_i)) - c \geq w(1 - F_u(s_i))$, i.e. the cost of skill acquisition

$$c \leq \beta(s) \equiv w(F_u(s) - F_q(s)), \tag{5}$$

the expected gain from doing so. Recalling that c has a distribution $G(c)$, the proportion of type i workers getting educated π_i equals the proportion of workers with cost less than $\beta(s_i)$, so that

$$\pi_i = G(\beta(s_i)), \; i = 1, W. \tag{6}$$

Given that $G(\beta(s))$ single peaked at $\tilde{s}$ and increasing if and only if $s < \tilde{s}$, the graph of $\pi_i(s_i)$ in the $s - \pi$ space, call it WW_i, $i = 1, W$, is inversely U-shaped.

Recalling that for a poor black worker of type c the cost of skill acquisition is

cm, we can mimic the preceding argument to write (denoting $\hat{\beta}(s) \equiv \frac{\beta(s)}{m}$)

$$\pi_2 = G(\hat{\beta}(s_2)). \tag{7}$$

Note that WW_W coincides with WW_1, and that WW_2 lies below WW_i, $i = W, 1$. Further, note that both curves peak at $\tilde{s}$.

As shall be clear shortly, this economy allows for multiple equilibria. For ease of exposition we impose assumption 2 below that helps us focus on the equilibria of interest.[18] To that end we define $\tilde{r}$ as solving

$$G(\hat{\beta}(\tilde{s})) = \frac{\phi(\tilde{s})}{r + \phi(\tilde{s})}.$$

Assumption 2. *(a)* $r > \tilde{r}$.

(b) There exists a unique $s''' > \tilde{s}$ *such that* EE *intersects* WW_2 *from above.*

(c) There exists no $s > \tilde{s}$ *such that* EE *intersects* WW_i, $i = 1, W$ *from above.*

Assumption 2(a) ensures that there is a unique $s' < \tilde{s}$ (resp. $s'' < \tilde{s}$) such that EE and WW_i, $i = 1, W$ (resp. EE and WW_2) intersect (see Figure 1). We then identify conditions such that a configuration $< \bar{s}_i, \bar{\pi}_i, \bar{\rho}_i >$, where $< \bar{s}_i, \bar{\pi}_i, \bar{\rho}_i > \equiv (\bar{s}_W, \bar{\pi}_W, \bar{\rho}_W; \bar{s}_1, \bar{\pi}_1, \bar{\rho}_1; \bar{s}_2, \bar{\pi}_2, \bar{\rho}_2)$ constitutes an equilibrium. Moreover, note that for all (π_j, s_j), where $s_j \in \{s', s'', s'''\}$ and (π_j, s_j) belongs to the graph of EE, we have that $\pi_j > 0$.

In the absence of affirmative action, a configuration $< \bar{s}_i, \bar{\pi}_i, \bar{\rho}_i >$, where $\bar{\rho}_i = \rho_i(\bar{s}_i, \bar{\pi}_i)$, constitutes an *equilibrium* if and only if, for all groups $i = 1, 2, W$, (a) given the cutoff $\bar{s}_i$, the proportion of group i workers acquiring the skill level is $\bar{\pi}_i$, and (b) given the level of skill acquisition $\bar{\pi}_i$, a cut-off of $\bar{s}_i$ maximises firm profits, i.e.

$$\pi_i = G(\beta(s(\pi_i))), \; i = 1, W, \tag{8}$$
$$\pi_2 = G((\hat{\beta}(s(\pi_2)))). \tag{9}$$

[18] In an earlier version of the paper we did not impose assumption 2. This does not change the analysis qualitatively, but makes some of the definitions substantially more involved, and consequently makes the analysis less transparent.

As in much of the theoretical literature on affirmative action, we shall be interested in two classes of equilibria, those with and without stereo-typing.

We say that an equilibrium does not involve stereo-typing if (a) the creamy layer blacks are held to the same standard as the white workers, and (b) whereas the poor black workers may be held to a higher standard, this arises solely from their relatively higher cost of acquiring education. Thus, given assumption 2 and Figure 1, one can say there is a unique stereo-typing equilibrium $< \bar{s}_i, \bar{\pi}_i, \bar{\rho}_i >$ that involves $\bar{s}_W = \bar{s}_1 = s'$, and $\bar{s}_2 = s''$.

An equilibrium is said to involve stereo-typing if the poor black workers are held to a standard that is 'too' high vis-a-vis the white workers. From assumption 2 and Figure 1, there is a unique stereo-typing equilibrium $< \bar{s}_i, \bar{\pi}_i, \bar{\rho}_i >$ that involves $\bar{s}_W = \bar{s}_1 = s'$, and $\bar{s}_2 = s'''$.[19,20]

(Figure 1 about here.)

Remark 1. *Given assumption 2, there is a unique stereo-typing, as well as a unique non-stereo-typing equilibrium. The preceding definitions would get more involved if assumption 2 is not imposed. Consider, for example, the stereo-typing equilibrium. If, say assumption 2(b) is not imposed, then there could be multiple stereo-typing equilibria and one would require an equilibrium selection mechanism among all such equilibria. One possibility is to impose the condition that the selected stereo-typing equilibrium satisfy, in addition, the property that there exists no other stereo-typing equilibrium with a lower cutoff for the poor black workers. Similar equilibrium selection issues would arise if assumption 2(c) is not satisfied. Finally if 2(a) is violated, then, given 2(c), there exists no* $s < \tilde{s}$ *such that* EE_i*,* $i = 1, W$ *intersects* WW *from above.*

3 Identity Based Affirmative Action

We then introduce the notion of identity based affirmative action formally and define equilibria in the presence of such affirmative action. We begin by identifying two political problems that may possibly afflict affirmative action policies. First, unlike in Coate and Loury (1993), we find that the removal of affirmative action can worsen

[19]Note that we focus on equilibria that are *locally stable* in that the absolute value of the slope of EE exceeds that of WW_i. For ease of exposition we shall refer to locally stable equilibria simply as equilibria.

[20]In an earlier version of the paper, that did not impose assumption 2, we allowed for the fact that there could be multiple equilibria with stereo-typing, focusing on equilibria that involved the 'least' stereo-typing as a selection device. This does not change the analysis qualitatively.

labour market outcomes for all poor black workers. Further, the interests of the whites and the creamy layer blacks may be opposed, in that while all creamy layer black workers prefer that affirmative action be adopted, the utility of all white workers decreases in that case. In Section 5 later, we compare the equilibrium outcome under such policies with that under class-based affirmative action.

Let $P(s_i', \pi_i)$ denote an employer's expected payoff from assigning a worker belonging to group i who emits a signal s_i' to task 1:

$$P(s_i', \pi_i) = \pi_i[1 - F_q(s_i')](x_q - w) - (1 - \pi_i)[1 - F_u(s_i')](x_u + w). \tag{10}$$

We consider an identity-based affirmative action policy that involves mandating 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.

$$\rho_W = \mu_1\rho_1 + \mu_2\rho_2, \tag{11}$$

where $\mu_i = \frac{\lambda_i}{\lambda_1+\lambda_2}$, $i = 1, 2$.

We then use (4) to define the functions

$$\hat{\rho}(s) \equiv \rho(s, G(\beta(s)), \text{ and } \hat{\hat{\rho}}(s) = \rho(s, G(\hat{\beta}(s))),$$

that will play an important role in the analysis. For $i = 1, W$, for example, $\hat{\rho}(s_i)$ denotes the fraction of group i workers that will be assigned to the skilled task when the employers adopt a cutoff of s_i for this group, and the workers respond optimally to this cutoff so that $G(\beta(s_i))$ workers in this group invest in skill acquisition. Note that $\hat{\rho}(0) = \hat{\hat{\rho}}(0) = 1$ and $\hat{\rho}(1) = \hat{\hat{\rho}}(1) = 0$.

Assumption 3. $\hat{\rho}(s)$ *and* $\hat{\hat{\rho}}(s)$ *are negatively sloped.*[21]

From Coate and Loury (1993) recall that in the absence of any income heterogeneity, given assumption 3, affirmative action completely resolves the problem of stereo-typing (in the sense that the resulting equilibrium not only involves no stereo-typing, it can moreover be sustained even after affirmative action is withdrawn). In

[21] In terms of model primitives, that $\hat{\rho}(s)$ is negatively sloped can equivalently be written as $\frac{f_u(s)-f_q(s)}{f_u(s)} < \frac{1}{G(\beta(s))+g(\beta(s))\beta(s)}$, $\forall s$, 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.

the present paper also this assumption plays a somewhat similar role, allowing us to abstract from issues of patronisation (to be defined in Section 5.1 later on). Given that Coate and Loury (1993) has studied the problem of patronisation extensively, to begin with we focus on equilibria without patronisation so that the various trade-offs can be analysed without any possible confounding effects arising out of patronisation, demonstrating that even in that case several results of interest emerge. In Section 5.1 later, we shall examine the outcome when assumption 3 is relaxed so that the equilibrium may involve patronisation.

We next turn to solving for equilibria in the presence of affirmative action. Because of affirmative action considerations, the employers' decisions in the various markets now become inter-linked, so that the constrained optimisation problem of an employer is:

$$\max_{s_1', s_2', s_W'} \sum_{i=W,1,2} \lambda_i P(s_i', \pi_i) + \gamma[\mu_1 \rho_1(s_1', \pi_1) + \mu_2 \rho_2(s_2', \pi_2) - \rho_W(s_W', \pi_W)], \tag{12}$$

where γ is the Lagrange multiplier on the affirmative action constraint. The first order condition with respect to s_i' generates the cutoff s_i as a function of π_i. Denoting this function by $EE_i(\gamma)$, $i = 1, 2, W$, we have

$$\frac{x_q - w - \gamma/\lambda_W}{x_u + w + \gamma/\lambda_W} = \frac{1-\pi_W}{\pi_W}\phi(s_W) \quad : \ EE_W(\gamma), \tag{13}$$

$$\frac{x_q - w + \gamma\mu_1/\lambda_1}{x_u + w - \gamma\mu_1/\lambda_1} = \frac{x_q - w + \gamma/\lambda_B}{x_u + w - \gamma/\lambda_B} = \frac{1-\pi_1}{\pi_1}\phi(s_1) \quad : \ EE_1(\gamma), \tag{14}$$

$$\frac{x_q - w + \gamma\mu_2/\lambda_2}{x_u + w - \gamma\mu_2/\lambda_2} = \frac{x_q - w + \gamma/\lambda_B}{x_u + w - \gamma/\lambda_B} = \frac{1-\pi_2}{\pi_2}\phi(s_2) \quad : \ EE_2(\gamma). \tag{15}$$

These conditions are intuitive, showing that as a result of affirmative action, employers act as if they have to pay a tax of $\frac{\gamma}{\lambda_W}$ on each white worker assigned to task one, while receiving subsidies of $\frac{\gamma}{\lambda_B}$ on each creamy layer and poor black worker they assign to task one.

Comparing with (3), where recall that (3) involves $r = \frac{1-\pi_i}{\pi_i}\phi(s_i)$, we have the following observation (see Figure 2):

Observation 1. *Let assumptions 1 and 2 hold. Fix* $\gamma > 0$.

(a) *For any* s_i*, the* π_i *solving (13), is greater than the* π_i *solving (3), so that graphically* $EE_W(\gamma)$ *lies to the right of* EE.

(b) *From (14) and (15),* $EE_1(\gamma)$ *and* $EE_2(\gamma)$ *coincide.*

(c) *For any* s_i*, the* π_i *solving (14) (as well as (15)) is less than the* π_i *solving* (3)*, so that* $EE_i(\gamma)$*,* $i = 1, 2$*, both lie to the left of* EE*.*

(Figure 2 about here.)

An outcome $< s_i^*, \pi_i^*, \rho_i^* > \equiv (s_W^*, \pi_W^*, \rho_W^*; s_1^*, \pi_1^*, \rho_1^*; s_2^*, \pi_2^*, \rho_2^*)$ constitutes an *equilibrium with affirmative action* if and only if, $\forall i = W, 1, 2$, (a) given the cutoff s_i^*, the proportion of group i workers acquiring the skill level is π_i^*, (b) given the level of skill acquisition π_i^*, a cut-off of s_i^* maximises firm profits, (c) $\rho_i^* = \rho_i(s_i^*, \pi_i^*)$, and (d) the affirmative action constraint (11) is satisfied. Formally, $< s_i^*, \pi_i^*, \rho_i^* >$ should satisfy

$$\pi_i = G(\beta(s(\pi_i))), i = 1, W, \quad (16)$$

$$\pi_2 = G(\widehat{\beta}(s(\pi_2))), \quad (17)$$

$$\rho_W = \mu_1\rho_1 + \mu_2\rho_2. \quad (18)$$

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_i from above, and moreover, for $i = 1, W$, $EE_i(\gamma^*)$ intersects WW_i from above exactly once, and $EE_2(\gamma^*)$ intersects WW_2 from above exactly twice.

Proposition 1 below establishes that there is an open set of parameter values such that an equilibrium with affirmative action exists. Given Proposition 1 below, we shall henceforth focus on equilibria with a positive Lagrange multiplier. The proof can be found in the appendix.

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

(a) There exists an open set of parameter values m *such that an affirmative action equilibrium with a positive Lagrange multiplier exists.*

(b) There can be at most two equilibria under affirmative action, one where $s_2^* < \tilde{s}$*, call it equilibrium without stereo-typing, and another where* $s_2^* > \tilde{s}$*, call it equilibrium with stereo-typing.*

We then establish some further properties of equilibria under affirmative action. The proof is in the Appendix.

Proposition 2. *Let assumptions 1, 2 and 3 hold. Denote the non-stereotyping equilibrium in the absence of affirmative action by* $< \bar{s}_i, \bar{\pi}_i, \bar{\rho}_i >$*. Under any affirmative action equilibrium* $< s_i^*, \pi_i^*, \rho_i^* >$ *one has that:*

(a) $\bar{s}_W < s_W^*,\ \bar{s}_1 > s_1^*,\ \bar{s}_2 > s_2^*,$

(b) $\bar{\rho}_W > \rho_W^*,\ \bar{\rho}_1 < \rho_1^*,\ \bar{\rho}_2 < \rho_2^*,$

(c) $\rho_1^* > \rho_W^* > \rho_2^*,$

(d) $s_2^* > s_W^* > s_1^*.$

3.1 Comparing Equilibria with and without Affirmative Action

We next compare the outcome under affirmative action with that in the absence of any affirmative action, in particular we examine how the imposition of affirmative action affects the utility of the workers belonging to the various groups. Proposition 3 below will play a critical role in answering this question. It demonstrates that while comparing across two different equilibria, it is sufficient to keep track of the cutoff signals across the two equilibria. Any group that faces a lower cutoff signal under one of these equilibria, strictly prefers that equilibrium to the other one.

Proposition 3. *Consider two distinct equilibria,* $< s_i', \pi_i', \rho_i' >$ *and* $< s_i'', \pi_i'', \rho_i'' >$*. Every individual in group* i *prefers* $< s_i', \pi_i', \rho_i' >$ *to* $< s_i'', \pi_i'', \rho_i'' >$ *if and only if* $s_i' < s_i''$*.*

Proof. For ease of exposition, let us denote $< s_i', \pi_i', \rho_i' >$ by E1, and $< s_i'', \pi_i'', \rho_i'' >$ by E2. We divide the individuals in group i into two classes.

(a) First consider group i 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 $s_i' < s_i''$, they have a greater chance of being assigned to task 1 under E1.

(b) Next consider individuals whose skill acquisition decision change across the two equilibria. Let $u_i(c, x, s, Ej)$ denote the utility of a group i individual with cost c under equilibrium Ej 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

$$u_i(c, N, s_i', E1) \geq u_i(c, Y, s_i', E1) \tag{19}$$
$$> u_i(c, Y, s_i'', E2), \tag{20}$$

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

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

$$u_i(c, Y, s_i', E1) \geq u_i(c, N, s_i', E1) \tag{21}$$
$$> u_i(c, N, s_i'', E2). \tag{22}$$

Thus all group i workers are strictly better off under E1 relative to E2. □

Note that the argument in Proposition 3 does not depend on whether there is any affirmative action policy in place or not, or on whether the affirmative action policy, if any, is identity, or class based. Thus we shall invoke this result while comparing across class and identity based affirmative action policies as well.

The next proposition suggests two political problems that may possibly afflict affirmative action policies. First, Coate and Loury (1993) argue that if assumption 3 holds and there is no heterogeneity in income, then all equilibria under affirmative action will lead to homogenous beliefs regarding the skill level of the various groups, so that removal of an affirmative action policy will not generate any change in the outcome. Consequently, temporary affirmative action policies can have a permanent effect (see their Proposition 2). Unlike Coate and Loury (1993) however, the removal of affirmative action can lead to a lower utility for all poor black workers in the current framework. This suggests that once enforced, affirmative action policies may be difficult to remove as doing so might lead to a significant loss of utility for the poor black workers, and consequently be politically unacceptable. Further, in Proposition 5(b) later, we shall argue that this is more likely to happen when the poor blacks are relatively 'large' in number. if so, then removing such affirmative action becomes even more difficult given that a significantly large section of the population is going to be adversely affected.

Second, we argue that with the implementation of affirmative action policies, while the creamy layer black workers are better off, white workers are worse off. Given that both these groups are likely to have significant voice in the political process, the implementation of such policies could be potentially problematic. Taken together, these two results demonstrate the importance of allowing for income heterogeneity within the target group.

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

(a) If the affirmative action equilibrium involves no stereo-typing, then all poor black workers will be worse off if affirmative action is removed.

(b) The utility of all creamy layer black workers is greater under affirmative action, while the utility of all white workers is lower.

Proof. Consider an equilibrium in the absence of affirmative action, $< \bar{s}_i, \bar{\pi}_i, \bar{\rho}_i >$. Note that under either the stereo-typing, or the non-stereo-typing equilibrium, $\bar{s}_W = \bar{s}_1 < \bar{s}_2$. Given assumption 2, we have that $\bar{\rho}_W = \bar{\rho}_1 > \bar{\rho}_2$, so that the affirmative action constraint (11) is not satisfied in either equilibrium.

(a) Note that the affirmative action constraint necessarily binds. Otherwise, the equilibrium will coincide with either the stereo-typing, or the non-stereotyping equilibrium, so that the affirmative action constraint cannot be satisfied. Given that the affirmative action constraint binds, we have that the Lagrange multiplier under affirmative action, call it γ^*, must be positive. Next, given that the equilibrium under affirmative action involves no stereo-typing, from Observation 1, it now follows that $s_2^* < \bar{s}_2 < \tilde{s}$, so that s_2 will increase once affirmative action is removed. The result then follows from Proposition 3.

(b) From the preceding argument the Lagrange multiplier under affirmative action must be positive. Given that $\gamma^* > 0$, from Observation 1 it then follows that $\bar{s}_1 > s_1^*$, and that $\bar{s}_W < s_W^*$. From Proposition 3 we have that the creamy layer blacks prefer that affirmative action happens, while the whites prefer that it does not.

□

Given Proposition 4(a), it is natural to ask if one can find conditions such that any equilibrium under affirmative action is without stereo-typing, so that removal of affirmative action policy necessarily leads to a decline in utility for the poor black workers. We need the following assumption before going further.

Assumption 4. *For all $\gamma \geq 0$, the maximal s such that EE_W intersects WW_W from above is less than s''''.*

Given this assumption, we find that the only equilibrium under affirmative action is non-stereo-typing whenever poor black workers form a significantly large fraction of all black workers. We recall from Proposition 4(a), in that case all poor black workers are necessarily worse off if affirmative action is removed. From Figure 1 recall that s'''' is the minimum s such that EE intersects WW_2 from below.

Proposition 5. *Let assumptions 1, 2, 3 and 4 hold. Fixing* λ_B*, there exists* $\widehat{\lambda}_2 \leq \lambda_B$*, such that for all* (λ_1, λ_2) *satisfying* $\lambda_B \geq \lambda_2 \geq \widehat{\lambda}_2$ *and* $\lambda_1 = \lambda_B - \lambda_2$*, there exists no affirmative action equilibrium with stereo-typing. In that case all poor black workers will be worse off if affirmative action is removed.*

Proof. We first prove that such an $\widehat{\lambda}_2$ exists. Suppose $\lambda_2 = \lambda_B$, and suppose to the contrary a stereo-typing equilibrium where $s_2^* > \tilde{s}$ exists. By Observation 1, $s_2^* > s''''$. Whereas by Observation 1 and assumption 4, $s_W^* < s''''$. Given assumption 3, the affirmative action constraint cannot be satisfied and no stereo-typing equilibrium exists. The existence of $\widehat{\lambda}_2$ then follows from continuity. Consequently, for all $\lambda_2 \geq \widehat{\lambda}_2$, all poor black workers will be worse off if affirmative action is removed. □

4 Class Based Affirmative Action

We next introduce class based affirmative action policy. As discussed earlier, class-based policies are being debated, and even adopted in many countries across the world in some sectors. In this section we define the notion of equilibrium under such a policy, before examining the implications of switching from an identity based to class based affirmative action policy in the following section.

A class based affirmative action policy mandates that the weighted average of the proportion of white *and* creamy layer black workers being assigned to task 1, equal that of the poor black workers. Defining $\lambda_1'' = \frac{\lambda_1}{\lambda_R}$ and $\lambda_W'' = \frac{\lambda_W}{\lambda_R}$, where $\lambda_R \equiv \lambda_1 + \lambda_W$, the class-based affirmative action condition is formally given by:

$$\rho_2 = \lambda_1''\rho_1 + \lambda_W''\rho_W. \tag{23}$$

Note that under a class based affirmative action policy not only are the white and creamy layer black workers at par as far as the cost of getting skilled is concerned, but the class based affirmative action policy also treats these two groups identically. Hence, in order to bring out the essential issues more sharply, we shall focus on symmetric equilibria where the cutoff signals and the extent of skill acquisition is the same across the two groups, i.e. $s_1 = s_W$, $\pi_1 = \pi_W$ and $\rho_W = \rho_1$. Thus the class based affirmative action condition can be re-written as:

$$\rho_2 = \rho_1 = \rho_W. \tag{24}$$

An outcome $< s_i^{**}, \pi_i^{**}, \rho_i^{**} >$, where $< s_i^{**}, \pi_i^{**}, \rho_i^{**} > \equiv (s_W^{**}, \pi_W^{**}, \rho_W^{**}; s_1^{**}, \pi_1^{**}, \rho_1^{**}; s_2^{**}, \pi_2^{**}, \rho_2^{**})$, constitutes an *equilibrium with class-based affirmative action* if and only if, $\forall i = W, 1, 2$, (a) given the cutoff s_i^{**}, the proportion of group i workers acquiring the skill level is π_i^{**}, (b) given the level of skill acquisition π_i^{**}, a cut-off of s_i^{**} maximises firm profits, (c) $\rho_i^{**} = \rho_i(s_i^{**}, \pi_i^{**})$, and (d) the class-based affirmative action constraint is satisfied, i.e. $\rho_2^{**} = \rho_1^{**} = \rho_W^{**}$.

Next, mimicking the earlier analysis, the optimisation problem of a firm yields the following first order conditions:

$$\frac{x_q - w - \gamma/\lambda_R}{x_u + w + \gamma/\lambda_R} = \frac{1-\pi_W^{**}}{\pi_W^{**}}\phi(s_W^{**}) \ : \ EE_W(\gamma), \tag{25}$$
$$\frac{x_q - w - \gamma/\lambda_R}{x_u + w + \gamma/\lambda_R} = \frac{1-\pi_1^{**}}{\pi_1^{**}}\phi(s_1^{**}) \ : \ EE_1(\gamma), \tag{26}$$
$$\frac{x_q - w + \gamma/\lambda_2}{x_u + w - \gamma/\lambda_2} = \frac{1-\pi_2^{**}}{\pi_2^{**}}\phi(s_2^{**}) \ : \ EE_2(\gamma). \tag{27}$$

For ease of exposition we continue to use the notations $EE_i(\gamma)$ for this case as well. Clearly under class-based affirmative action, $EE_1(\gamma)$ and $EE_W(\gamma)$ coincide, and both lie above EE, whereas $EE_2(\gamma)$ lies below EE. As in case of identity-based affirmative action, we shall restrict attention to equilibria that are locally stable, and moreover, for $i = 1, W$, $EE_i(\gamma^{**})$ intersects WW_i from above exactly once, and $EE_2(\gamma^{**})$ intersects WW_2 from above exactly twice.

Proposition 6. *Let assumptions 1, 2 and 3 hold.*

(a) There exists an open set of m such that a class-based affirmative action equilibrium with a positive Lagrange multiplier exists.

(b) There can be at most two equilibria under class based affirmative action, one of which involves $s_2^{**} < \tilde{s}$, which we call a non-stereo-typing equilibrium, and another which involves $s_2^{**} > \tilde{s}$, which we call a stereo-typing equilibrium.

(c) Under any equilibria the poor blacks face a lower cutoff compared to the other groups, i.e. $s_2^{**} < s_1^{**} = s_W^{**}$.

Proof. (a) Given continuity, the argument follows from Proposition 1 in Coate and Loury (1993).

(b) One can mimic the argument in Proposition 1(b) to argue that there can be at most two equilibria, one where $s_2^{**} < \tilde{s}$, and another where $s_2^{**} > \tilde{s}$.

(c) We note the affirmative constraint ensures that $\rho_1^{**} = \rho_2^{**} = \rho_W^{**}$. The claim now follows since, $\forall s$, $\hat{\hat{\rho}}(s) < \hat{\rho}(s)$, and from assumption 3 both $\hat{\hat{\rho}}(s)$ and $\hat{\rho}(s)$ are negatively sloped. □

We then argue that in the presence of assumption 4, there is a unique equilibrium that does not involve stereo-typing, and moreover, in that case, all poor black workers will be worse off if affirmative action is removed.

Proposition 7. *Let assumptions 1, 2, 3 and 4 hold.*

(a) There can be at most one equilibrium under class based affirmative action, one that involves no stereo-typing equilibrium.

(b) While all poor black workers will be worse off if class-based affirmative action is removed, both the white and the creamy layer black workers will be better off.

Proof. (a) Suppose to the contrary there is an equilibrium where $s_2^{**} > \tilde{s}$. Then, given assumption 4, it must be that $s_2^{**} > s_1^{**} = s_W^{**}$, so that the affirmative action constraint cannot be satisfied.

(b) Given that the class-based affirmative action constraint binds, it must be that the associated Lagrange multiplier is strictly positive. Consequently from Proposition 7(a) and Observation 1 it follows that $s_2^{**} < s' < s_1^{**} = s_W^{**}$. The result follows since recall that in the non-stereo-typing equilibrium, denote it by $< \bar{s}_i, \bar{\pi}_i, \bar{\rho}_i >$, $\bar{s}_W = \bar{s}_1 = s'$ and $\bar{s}_2 = s''$, where recall that $s'' > s'$. □

Taking Propositions 5(b) and 7(b) together, we find that irrespective of whether affirmative action is identity or class-based, removing affirmative action may lead to a loss in utility for the poor blacks. This in turn may make removal of any such policy politically difficult, especially if poor black workers form a significantly large section of the black population.

5 Comparing Identity and Class Based Affirmative Action

Next we consider the implications of switching from an identity-based to a class-based affirmative action policy. Proposition 8 below demonstrates that, relative to identity based affirmative action, all creamy layer black workers are worse off under class based affirmative action. Moreover, all poor black workers are better off, whereas the white workers are worse off under a class-based policy.

We need the following regularity condition before proceeding further. We later discuss the implications of relaxing this assumption.

Assumption 5. $\frac{1-G(\beta(s))}{G(\beta(s))}\phi(s)$ *and* $\frac{1-G(\hat{\beta}(s))}{G(\hat{\beta}(s))}\phi(s)$ *are both decreasing in* s *for all* s.[22]

The next result compares identity based affirmative action policy with a class-based one.

Proposition 8. *Let assumptions 1, 2, 3 and 5 hold. All poor black workers are better off, whereas all white, as well as all creamy layer black workers are worse off in the event of regime switch from identity to class based affirmative action.*

Proof. Let $< s_i^*, \pi_i^*, \rho_i^* >$ denote an equilibrium under identity based affirmative action, while $< s_i^{**}, \pi_i^{**}, \rho_i^{**} >$ denotes an equilibrium under class-based affirmative action.

To begin with note that $s_1^{**} > s_1^*$ (this follows as $EE_1(\gamma)$ shifts down under identity-based affirmative action, and shifts up under class based affirmative action). This implies that $\rho_1^* > \rho_1^{**}$ (from assumption 3).

We next argue that $\rho_2^{**} > \rho_2^*$. Suppose to the contrary that $\rho_2^{**} \leq \rho_2^*$. This implies that $s_2^{**} \geq s_2^*$ (from assumption 3). Given (15), (27) and 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 then implies that $s_W^{**} \leq s_W^*$ (from (13), (25) and the fact that $\frac{1-G(\beta(s))}{G(\beta(s))}\phi(s)$ is decreasing in s), and hence $\rho_W^{**} \geq \rho_W^*$ (from assumption 3).

Next, given that $\rho_2^{**} = \rho_1^{**} = \rho_W^{**}$, we have that

$$\rho_W^{**} = \mu_1\rho_1^{**} + \mu_2\rho_2^{**}.$$

Thus,

$$\rho_W^* \leq \rho_W^{**} = \mu_1\rho_1^{**} + \mu_2\rho_2^{**} < \mu_1\rho_1^* + \mu_2\rho_2^*,$$

which contradicts the affirmative action constraint under identity-based affirmative action. Hence, $\rho_2^{**} > \rho_2^*$, and thus $s_2^{**} < s_2^*$ (from assumption 3). This in turn implies that $\gamma^* < \gamma^{**}$ (mimicking our earlier argument), so that from Observation 1

[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. If, for example, (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 $s \in [0, 1]$, then it is straightforward to check that this condition holds.

and assumption 3, $\rho_W^{**} < \rho_W^*$. Thus from assumption 2, $s_W^{**} > s_W^*$. The result now follows from Proposition 3. $\square$

Intuitively, under class based affirmative action, the creamy layer black workers no longer belong to the target group, unlike under identity based affirmative action. Consequently, these workers face relatively higher standards and are worse off under class based affirmative action. This implies that the poor blacks would be better off as well, since otherwise the shadow price of equality must be lower when affirmative action is based on income, rather than on identity. A lower shadow price of equality, in turn, would imply that the standards for the whites would be higher under identity-based affirmative action, so that they would be worse off. This, however, violates the identity-based affirmative action constraint, as both creamy layer and poor blacks cannot be better off under identity-based affirmative action, while the whites are worse off. Thus, class-based affirmative action increases the shadow price of equality, making poor blacks better off, and whites worse off.

How restrictive is assumption 5? We next argue that this assumption is automatically satisfied whenever poor black workers constitute a significant fraction of all black workers and assumption 4 holds. 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, the preceding proposition goes through whenever all equilibria involves cutoff signals that are less than $\tilde{s}$. From Proposition 5 earlier, we recall that for any $\lambda_2 \geq \hat{\lambda}_2$, it is the case that any equilibrium under identity-based affirmative action is a non-stereo-typing equilibrium. Further, from Proposition 7 earlier, recall that any equilibrium under class-based affirmative action must be non-stereo-typing. Thus, for all $\lambda_2 \geq \hat{\lambda}_2$, all equilibria under either form of affirmative action are non-stereo-typing, so 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 decreasing over the relevant range and Proposition 8 goes through.

Proposition 8 suggests that a move from identity based to class based affirmative action may face political hurdles, as both the whites, as well as creamy layer blacks, groups with significant voice would be worse off as a result of such a move. Moreover, while white workers are worse off under both types of affirmative action, they may be relatively more willing to support identity-based, rather than class-based affirmative action policies. This is interesting given that most countries start with identity based affirmative action.

5.1 Patronizing Equilibria

Finally we introduce the notion of patronization first developed in Coate and Loury (1993). We find that even under a patronizing equilibrium, the creamy layer black workers are worse off in case of a move from identity to class-based affirmative action, and the interests of the whites and the creamy layer blacks are going to be opposed. To that end we examine an economy where $\hat{\rho}(s)$ and and $\hat{\hat{\rho}}(s)$ are non-monotonic.

Assumption 6. *Let* $\hat{\rho}(s)$ *and* $\hat{\hat{\rho}}(s)$ *be both increasing over some interval* $(\underline{s}, \bar{s})$, *where* $0 < \underline{s} < \bar{s} < 1$.

We next define the notion of a patronizing equilibrium under identity based affirmative action, as one where firms have correct beliefs about the inferiority of the poor blacks and therefore use a lower standard to ensure that poor blacks are assigned to task 1 at a large enough rate to ensure that the affirmative action constraint is being met.

An equilibrium $< s_i^*, \pi_i^*, \rho_i^* >$ under identity based affirmative action is said to be *patronizing* if and only if $s_2^* < s_W^*$.

An equilibrium $< s_i^{**}, \pi_i^{**}, \rho_i^{**} >$ under class based affirmative action is said to be *patronizing* if and only if $s_2^{**} < s_i^{**},\ i = 1, W$.

Note that for $m = 1$, our framework coincides with Coate and Loury (1993). Consequently, from continuity, Proposition 4 in Coate and Loury (1993) guarantees the existence of a patronising equilibrium under both identity as well as class based affirmative action for an open set of parameter values (that requires λ_W to be sufficiently large) for m sufficiently close to 1.

In the next proposition we examine the effects of a switch from identity-based to class-based affirmative action in the presence of patronization. Interestingly, under a patronising equilibrium we find that interests of the poor black and the white workers are necessarily opposed as far as labour market outcomes are concerned, though the effects on these two groups, relative to the affirmative action policy, are ambiguous. Consequently, class-based affirmative action is unlikely to have much political backing whenever the poor black workers actually gain in the labour market, as both the whites and creamy layer blacks would then be worse off.

Proposition 9. *Let assumptions 1, 2, 5 and 6 hold, and suppose that a patronising equilibrium exists under both identity based, as well as class based affirmation action.*

(a) The creamy layer black workers are worse off under class based affirmative action.

(b) The interests of the whites and the poor black workers are opposed in the sense that white workers are better off under class based affirmative action if and only if the poor black workers are worse off.

Proof. Mimicing the argument in Proposition 8, note that $s_1^{**} > s_1^*$. Further, if $\gamma^{**} \geq \gamma^*$, then $s_2^{**} \leq s_2^*$ and $s_W^{**} \geq s_W^*$ from Observation 1. Whereas if $\gamma^{**} < \gamma^*$, then $s_2^{**} > s_2^*$ and $s_W^{**} < s_W^*$. □

Intuitively, the reason for creamy layer blacks being worse off under class-based affirmative action is the same as for Proposition 8. Turning to the poor black and white workers, if the poor blacks are worse off (respectively better-off) under class-based affirmative action, then the shadow price of equality must be lower (respectively higher), so that the whites must be better off (respectively worse off).

6 Conclusion

We examine affirmative action in a framework with statistical discrimination, as well as income heterogeneity among the beneficiaries of affirmative action. We then use this framework to examine a set of questions that are of importance to the countries practicing affirmative action.

We find that attempts to roll back affirmative action policies can be problematic. The presence of income heterogeneity means that - unlike in Coate and Loury (1993) - removal of affirmative action may worsen labour market outcomes for poor blacks, so that temporary affirmative action may not have a permanent effect. This finding suggests that once affirmative action is in place, it might be politically very difficult to remove such policies. A case in point could be India, where all major political parties are careful to support the continuation of affirmative action policies.[23] Relatedly one finds that in the US, as universities in many states began to limit their use of race-based affirmative action policies, minority enrolment in these universities dropped, and segregation increased, which is consistent with our results.[24]

We then examine another politically charged issue - how moving to class-based affirmative action affects the three groups. Our findings indicate that such a transition

[23] Affirmative action, which was supposed to be phased out within 10 years, is now continuing for over 60 years in India.

[24] http://www.nyu.edu/classes/jackson/social.issues/papers/AfActGrG.html.

could be politically divisive, with at least two of the three groups, which necessarily includes the creamy layer blacks, losing out in the labour market as a result. If there is no patronization in equilibrium, then such a transition would make the poor blacks better off, but hurt both the whites and the creamy layer blacks in the labour market.

Taken together, these results suggest that affirmative action is politically an extremely sensitive issue, particularly in the presence of income heterogeneity. Given the complexities of the trade-offs involved, implementation is going to be as much as a political challenge as an economic one, and will require serious and sustained efforts at consensus building.

7 Appendix

Proof of Proposition 1. (a) Define $\hat{s}_i$ as solving

$$\hat{s}_i(\gamma) = \{\min s_i | s_i \text{ lies at the intersection of } EE_i(\gamma) \text{ and } WW_i\}.$$

Note that $\hat{s}_1(0) = \hat{s}_W(0)(= \bar{s}_1 = \bar{s}_W)$. Further, given $r > \tilde{r}$, it is the case that $\hat{s}_i(0) < \tilde{s}$, $i = 1, W$. Moreover, from assumption 2, $\hat{s}_2(0) = \bar{s}_2 < \tilde{s}$. Note that $\bar{s}_1 = \bar{s}_W < \bar{s}_2$, so that from 3 one has that $\hat{\rho}(\bar{s}_1) > \hat{\rho}(\bar{s}_2)$. Consider any $r > \tilde{r}$, and m such that $1 < m < \tilde{m}(r)$. To show that there exists $\gamma^*(r, m) > 0$, such that $\hat{s}_i(\gamma^*)$ satisfies affirmative action.

Let γ increase from zero. Note that $\hat{s}_i(\gamma)$ is continuous, with $\hat{s}_1(\gamma), \hat{s}_2(\gamma)$ being decreasing, and $\hat{s}_W(\gamma)$ being increasing in γ. Next define

$$D(\gamma) = \mu_1\hat{\rho}(\hat{s}_1(\gamma)) + \mu_2\hat{\rho}(\hat{s}_2(\gamma)) - \hat{\rho}(\hat{s}_W(\gamma)).$$

Note that

$$\begin{aligned}
D(0) &= \mu_1\hat{\rho}(\hat{s}_1(0)) + \mu_2\hat{\rho}(\hat{s}_2(0)) - \hat{\rho}(\hat{s}_W(0)) \\
&= \mu_1\hat{\rho}(\bar{s}_1) + \mu_2\hat{\rho}(\bar{s}_2) - \hat{\rho}(\bar{s}_W) \\
&= \mu_2\hat{\rho}(\bar{s}_2) - (1 - \mu_1)\hat{\rho}(\bar{s}_1) \text{ (since } \bar{s}_1 = \bar{s}_W) \\
&= \mu_2(\hat{\rho}(\bar{s}_2) - \hat{\rho}(\bar{s}_1)) < 0,
\end{aligned}$$

where the last inequality follows from assumption 3 and the fact that $\bar{s}_1 = \bar{s}_W < \bar{s}_2$. Further, as $\gamma \to \lambda_B(x_u + w)$, for $i = 1, 2$, $\hat{s}_i(\gamma)$ goes to zero (from 1), so that $\hat{\rho}(s_i)$ goes to 1, and consequently $[\mu_1\hat{\rho}(\hat{s}_1(\gamma)) + \mu_2\hat{\rho}(\hat{s}_2(\gamma))]|_{\gamma \to \lambda_B(x_u+w)} = 1$. Given that $\hat{\rho}(\hat{s}_W(\lambda_B(x_u + w)) < 1$, 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 $D(\gamma^*) = 0$.

(b) Consider an affirmative action equilibrium $< s_i^*, \pi_i^*, \rho_i^* >$, with an associated Lagrange multiplier of γ^*. Note that the affirmative action constraint implies that in this equilibrium $\rho_W^* = \mu_1 \rho_1^* + \mu_2 \rho_2^*$. Any other affirmative action equilibrium must involve a different γ', where without loss of generality let $\gamma' > \gamma^*$. Then from Observation 1, the affirmative action constraint cannot be satisfied. $\square$

Proof of Proposition 2. (a) Note that Proposition 2(a) follows from Observation 1 (see Figure 2).

(b) Given Proposition 2(a), Proposition 2(b) follows from the fact that $\hat{\rho}(s)$ is negatively sloped (from 2).

(c) and (d) Observe that from Proposition 2(a), $s_W^* > \bar{s}_W = \bar{s}_1 > s_1^*$. Hence given assumption 2, $\hat{\rho}(s_1^*) > \hat{\rho}(s_W^*)$. Further given that (a) $\hat{\rho}(s_1^*) > \hat{\rho}(s_W^*)$, and (b) from the affirmative action constraint, $\hat{\rho}(s_W^*)$ equals the average of $\hat{\rho}(s_1^*)$ and $\hat{\rho}(s_2^*)$, we have that $\hat{\rho}(s_W^*) > \hat{\rho}(s_2^*)$. Next, given $\hat{\rho}(s)$ is decreasing, $s_2^* > s_W^* > s_1^*$. $\square$

8 Bibliography

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

Feagin, J.R., and Porter, A (1995) : Affirmative action and African Americans: rhetoric and practice. Humboldt Journal of Social Relations, 21(2): 81-103.

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.

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

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

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

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

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

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

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.

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

Figure 1

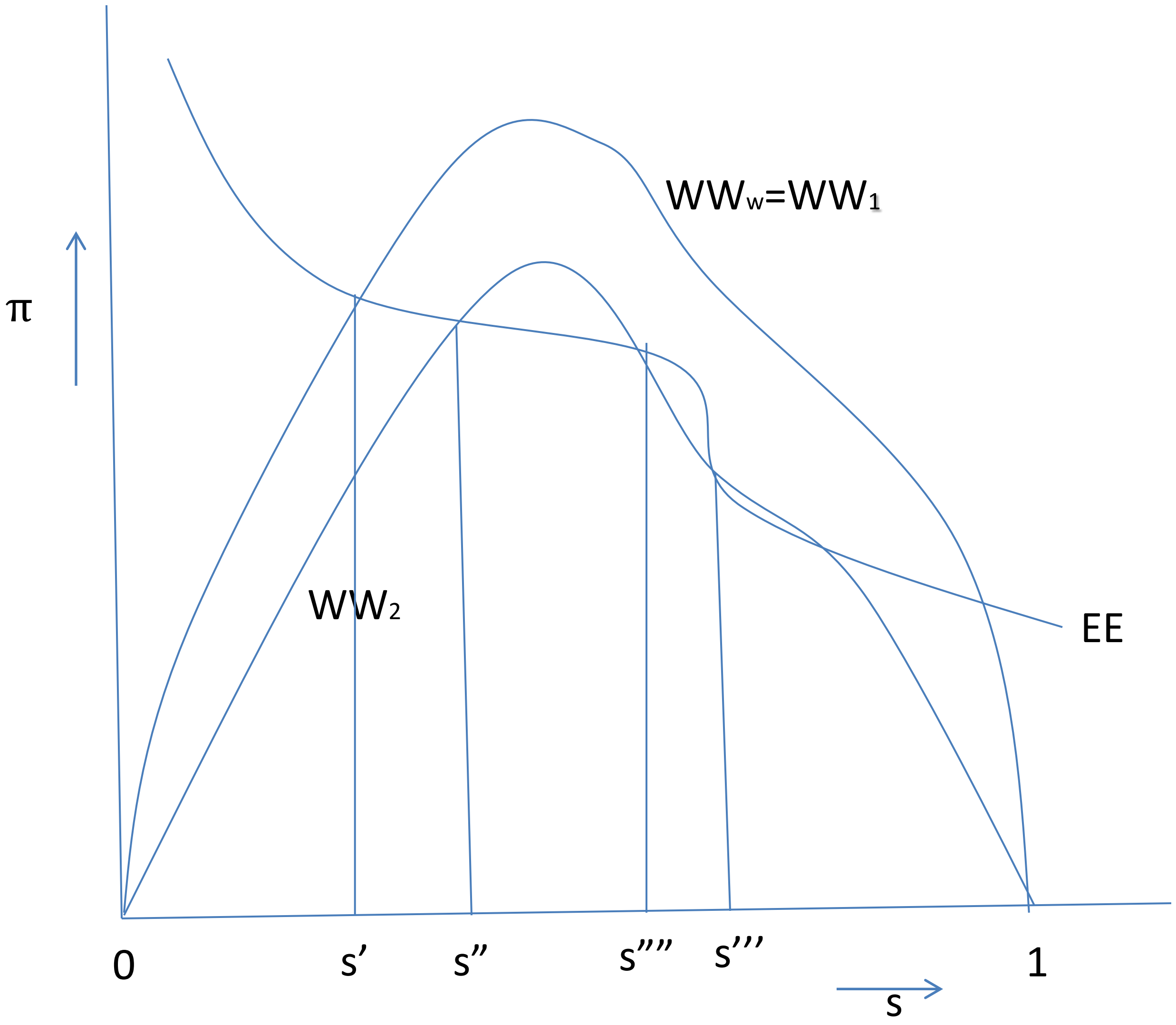

Figure 2