A Remark on Color-Blind Affirmative Action

By Debraj Ray and Rajiv Sethi[1]

April 2009

Elite colleges and universities in the United States have recently faced a number of legal challenges that restrict their use of explicitly race-contingent admissions policies. Since these institutions continue to seek broad representation from different social groups (and to view campus diversity as an essential ingredient in the provision of a first-rate education), they have strong incentives to adjust their admissions criteria in order to attain diversity goals through less direct means. There is considerable evidence that this process is well underway, and a literature dealing with the efficiency implications of color-blind affirmative action policies has emerged (Chan and Eyster 2003, Fryer and Loury 2007, Fryer et al. 2008, Epple et al. 2008).

This note is concerned with one possible feature of a color-blind policy: the use of admissions criteria that, while uniform across all social groups, are *non-monotone* in measures of past performance (such as high school grades or standardized test scores). Under such polices individuals with lower scores may receive admission with greater likelihood than those in the same social group with somewhat higher scores, simply because their scores fall into a range which is disproportionately populated by members of an underrepresented group. The possibility that optimal color-blind policies might have this structure was recognized by Chan and Eyster (2003), although they assumed for the purposes of their analysis that only monotone rules were feasible. More recently, Epple et al. (2008) have computed color-blind admissions policies using a calibrated general equilibrium model, and found them to be non-monotonic in scores (conditional on other nonracial characteristics such as income).

We establish here that under weak conditions that apply generically, non-monotonicity is not simply a theoretical possibility or the predicted outcome under a plausible calibration, but a *necessary* property of score-maximizing color-blind admissions policies. In addition, we argue that blind rules can generate significantly greater disparities in mean scores across groups conditional on acceptance than would arise if explicitly race-contingent policies were permitted. This is most easily seen in the case of score-maximizing (and hence non-monotone) policies, but applies also to blind policies that are constrained to be monotone. We also briefly discuss the manner in which non-monotone policies can be implemented in practice, given some natural incentive compatibility issues that arise from their use.

[1]New York University and Barnard College, Columbia University. We thank (without implicating) Dennis Epple and Glenn Loury for comments on an earlier version.

Consider a population of measure 1 composed of two groups (black and white) with population shares β and $1-\beta$ respectively. The within-group test score distribution functions are $F_b(\theta)$ and $F_w(\theta)$ respectively, and are assumed to possess corresponding continuous densities $f_b(\theta)$ and $f_w(\theta)$. The aggregate distribution function is

$$F(\theta) = \beta F_b(\theta) + (1-\beta)F_w(\theta),$$

with corresponding density

$$f(\theta) = \beta f_b(\theta) + (1-\beta)f_w(\theta).$$

We assume that whites score higher as a group relative to blacks in the sense of first order stochastic dominance:

$$F_b(\theta) \geq F_w(\theta) \tag{1}$$

for all θ, with strict inequality whenever $F_b(\theta) > 0$ and $F_w(\theta) < 1$.

Only a proportion α of the total population can receive admission. Define θ^* as the test score cutoff if admission is based on scores alone, with the highest scorers admitted. Then

$$1 - F(\theta^*) = \alpha.$$

Let β^* denote the share of the admitted population that is from the disadvantaged group under this policy. Then

$$\alpha\beta^* = \beta(1 - F_b(\theta^*)).$$

It follows from (1) that $\beta^* < \beta$ (members of the disadvantaged group are underrepresented in the population of admitted students). Among those accepted, the mean score for students belonging to group $i \in \{b, w\}$ is

$$m_i = \frac{1}{1 - F_i(\theta^*)} \int_{\theta^*}^{\infty} \theta f_i(\theta) d\theta,$$

and the mean score among all accepted students is $\beta m_b + (1-\beta)m_w$.

Now suppose that a target level of representation $\tilde{\beta} > \beta^*$ is desired, and the institution wishes to maximize the average score among admitted students subject to the constraint that this target, as well as the capacity constraint, are both met. If race-contingent policies are permissible, this may be accomplished by selecting a distinct score threshold for each group and admitting all those whose scores exceed the threshold corresponding to group to which they belong. Let θ_b and θ_w denote these thresholds. Then the diversity constraint is

$$\beta(1 - F_b(\theta_b)) = \tilde{\beta}\alpha,$$

and the capacity constraint is

$$1 - \beta F_b(\theta_b) - (1-\beta)F_w(\theta_w) = \alpha.$$

There is a unique value of θ_b consistent with the diversity constraint and, given this, a unique value of θ_w consistent with the capacity constraint. It is clear that for any target $\tilde{\beta} > \beta^*$, we must have $\theta_w > \theta^* > \theta_b$; the advantaged group will face a more demanding threshold for admission. The mean score for accepted students belonging to group $i \in \{b, w\}$ with this (sighted) affirmative action policy is

$$m_i^s = \frac{1}{1 - F_i(\theta_i)} \int_{\theta_i}^{\infty} \theta f_i(\theta) d\theta,$$

and the mean for the entire population of admitted students is $\beta m_b^s + (1 - \beta) m_w^s$.

If colleges are prevented from making explicit use of group identity, they must apply the same set of admissions criteria to members of both groups. This makes it costlier (in terms of the mean score among accepted students) but not impossible to meet a diversity target. Following Fryer and Loury (2007), we refer to an admissions policy as *color-blind* if a student's likelihood of acceptance under that policy depends only on his score and not on his identity, and *sighted* if it is explicitly race-contingent. Any color-blind policy can be represented by a function $p(\theta)$, interpreted as the probability of acceptance for someone having score θ. We say that a color blind admissions policy is *monotone* if $p(\theta)$ is nondecreasing in θ.

The policy of simply admitting the highest scorers may then be written as

$$p(\theta) = \begin{cases} 1 & \text{if } \theta \geq \theta^* \\ 0 & \text{otherwise} \end{cases}$$

This is clearly a monotone policy, which results in the default black population share β^* among admitted students. Any target $\tilde{\beta} \in [\beta^*, \beta]$ can be met using a policy that is color blind and monotone, for instance by the appropriate choice of two thresholds such that those above the higher threshold are admitted with certainty, those below the lower one rejected, and those between thresholds admitted with a suitably chosen probability. Chan and Eyster show that such a 'two-step' policy can be used to maximize mean entering scores within the class of rules that are both blind and monotone.

A policy is *deterministic* if $p(\theta) \in \{0, 1\}$ for (almost) all θ in the support of test scores. The score-maximizing policy with no affirmative action is a deterministic policy. A little thought shows that a policy that needs to meet a target $\tilde{\beta} > \beta^*$ cannot be both deterministic and monotone.

Say that $p(\theta)$ is a *score-maximizing color-blind policy* (*with target* $\tilde{\beta}$) if it is a solution to the following problem (with limits of integration omitted):

$$\max_{p(\theta)} \int \theta p(\theta) f(\theta) d\theta \tag{2}$$

subject to

$$\int p(\theta) f(\theta) d\theta = \alpha, \tag{3}$$

and

$$\beta \int p(\theta) f_b(\theta) d\theta \geq \tilde{\beta}\alpha. \tag{4}$$

We impose the following condition on the distributions of test scores:

[G] For any θ such that $f_b(\theta) > 0$, $f_b(\theta)/f(\theta)$ is not locally affine in θ.[2]

This condition is extremely mild and simply rules out degenerate cases in which the ratio of black and white score densities moves over some interval in a way that is precisely linear in the score.

Proposition **1.** *Under Condition* [G]*, if $p(\theta)$ is a score-maximizing color-blind policy with target $\tilde{\beta} > \beta^*$, then it is deterministic.*

In particular, such a policy cannot be monotone.

Proof. Consider the problem described in (2)–(4). Standard arguments for infinite-dimensional convex programming (see, e.g. Rockafellar (1974), especially Example 1, pp. 7 and 18, as well as Example 1′, p. 23) allow us to assert the existence of multipliers λ and μ for the constraints (3) and (4), such that a score-maximizing policy must maximize

$$\int p(\theta) \left[\theta f(\theta) + \lambda f(\theta) + \mu\beta f_b(\theta)\right] d\theta, \tag{5}$$

subject to the constraint that for every θ, $p(\theta) \in [0, 1]$.

The expression (5) tells us that the maximizing policy must be deterministic provided

$$\theta f(\theta) + \lambda f(\theta) + \mu\beta f_b(\theta) \neq 0$$

for almost every θ in the support of test scores. Because $f(\theta) > 0$ a.e. on the support of test scores, this condition is equivalent to

$$(\theta + \lambda) + \mu\beta \frac{f_b(\theta)}{f(\theta)} \neq 0$$

for almost every θ such that $f(\theta) > 0$. But this follows right away from the continuity of f_b and f, and the genericity condition [G].

In particular, then, p cannot be monotone. For if it were monotone and deterministic, it is easy to check that either condition (3) or (4) must be violated whenever $\tilde{\beta} > \beta^*$. ■

[2] More precisely, there is no interval around θ such that for all θ' in that interval, $[f_b(\theta)/f(\theta)] = A + B\theta$ for constants A and B.

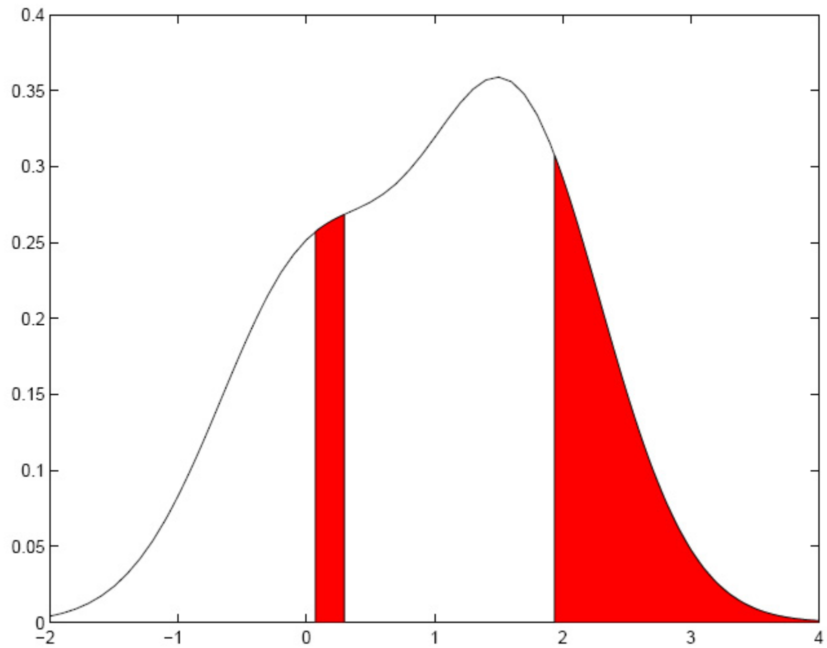

Figure 1. A Score-Maximizing Color-Blind Policy

The following numerical example illustrates the structure of a score-maximizing admissions policy. Suppose that both $f_b(\theta)$ and $f_w(\theta)$ are normal with variance $1/2$ and means μ_b and μ_w respectively. Then for $i \in \{b, w\}$,

$$f_i(\theta) = \frac{1}{\sqrt{\pi}} e^{-(\theta-\mu_i)^2}$$

and

$$F_i(\theta) = \frac{1}{2}(1 + \mathrm{erf}(\theta - \mu_i)).$$

Suppose that

$$\alpha = \frac{1}{4},\ \beta = \frac{2}{5},\ \tilde{\beta} = \frac{1}{5},\ \mu_b = 0,\ \mu_w = \frac{8}{5}.$$

Then the score-maximizing blind affirmative action policy is

$$p(\theta) = \begin{cases} 1 & \text{if } \theta \in (\theta_1, \theta_2) \cup [\theta_3, \infty) \\ 0 & \text{otherwise} \end{cases} .$$

where $(\theta_1, \theta_2) = (0.07, 0.30)$ and $\theta_3 = 1.94$. This policy is illustrated in Figure 1.

It is instructive to examine in detail the admission requirements and mean scores by group for admitted students under four policy alternatives: no affirmative action, the sighted policy, the blind score-maximizing policy, and the two-step blind and monotone policy. The following table summarizes admissions criteria and mean scores under each of these.

AA Policy	Admission requirement	$\tilde{\beta}$	m	m_b	m_w
None	$\theta > 1.76$	0.01	2.26	1.99	2.27
Sighted	$\theta_b > 0.81$, $\theta_w > 1.90$	0.20	2.13	1.16	2.37
Blind score-maximizing	$\theta \in (0.07, 0.30) \cup (1.94, \infty)$	0.20	1.87	0.23	2.28
Blind monotone	$\theta > 2.13$, $\theta \in (-0.92, 2.13)$ with prob 0.14	0.20	1.75	0.15	2.15

The mean score among all admitted students is highest when no affirmative action policy is implemented, but the level of diversity falls well below the desired threshold. Sighted affirmative action allows the diversity target to be met at some cost in terms of the overall mean score among admitted students. Relative to the case of no affirmative action, admitted white students have higher mean scores, and admitted black students have lower mean scores. This is a necessary consequence of the policy and does not depend on the particular specification used here.

The blind score-maximizing policy results in a much greater disparity in mean scores across the two groups when compared with the sighted policy. The reasons for this are apparent from Figure 1: the diversity constraint is met by recruiting students from a part of the overall score distribution that is heavily populated by the underrepresented group, but which falls some distance to the left of the cutoff point for remaining students. A wide gap between mean entering scores across social groups also arises in the case of blind policies that are constrained to be monotone. The reasons are similar: in the case of monotone policies the diversity constraint is met be accepting students with low probability across a very broad range of the score distribution. In addition, the monotone policy has lower scores not only overall but also within each group when compared with the score-maximizing admission rule.

How might non-monotone policies be implemented in practice? One possibility is to focus recruitment efforts on two disjoint applicant pools: those from elite high schools with high levels of past performance on standardized tests, and those from largely segregated schools with lower levels of past performance which allow diversity goals to be met in a manner that is not formally contingent on applicant identity. The former pool helps raise the value of the objective function, while the latter pool allows the diversity constraint to be met at relatively little cost in terms of overall mean scores. As long as movement of students across such disjoint pools is limited, such polices need not violate incentive compatability constraints. Note, however, that admissions policies must be monotone *conditional on social location* if they are to be incentive compatible (Loury, personal communication). Taking explicit account of recruitment efforts with multiple applicant pools and limited mobility is beyond the scope of this note but would seem to be a worthwhile exercise.

A non-monotone policy has the property that *within each group* some students with lower scores are admitted while others with higher scores are denied. As noted by Chan and Eyster, this violates certain intuitive notions of fairness. Furthermore, blind policies (both monotone and score-maximizing) can widen the disparity between black and white scores conditional on admission, resulting in the reinforcement and entrenchment of negative stereotypes. As Epple et al. (2008, p. 476) note, a common justification for affirmative action is that "racial diversity in student bodies promotes cross-racial understanding and breaks down stereotypes, which better prepares students for an increasingly diverse workplace." This particular goal is undermined by the use of color-blind policies to the extent that they induce larger gaps between groups in mean scores conditional on acceptance than would arise under the more traditional forms of affirmative action.

References

[1] Chan, Jimmy, and Erik Eyster (2003). "Does Banning Affirmative Action Lower College Student Quality?" *American Economic Review* 93: 858-872.

[2] Epple, Dennis, Richard Romano and Holger Sieg (2008). "Diversity and Affirmative Action in Higher Education." *Journal of Public Economic Theory* 10: 475-501.

[3] Fryer, Roland G. and Glenn C. Loury (2007). "Valuing Identity: The Simple Economics of Affirmative Action." Unpublished manuscript, Brown University.

[4] Fryer, Roland G., Glenn C. Loury, and Tolga Yuret (2008). "An Economic Analysis of Color-Blind Affirmative Action." *Journal of Law, Economics, and Organization* 24: 319-355.

[5] Rockafellar, R. Tyrone (1974). *Conjugate Duality and Optimization*, Philadelphia, PA: SIAM.