Color-Blind Affirmative Action

Roland G. Fryer, Jr., Glenn C. Loury, and Tolga Yuret*

October 2003

Abstract

This paper presents a conceptual framework for understanding the consequences of the widespread adoption of "race-neutral alternatives" to conventional racial affirmative action policies in college admissions. A simple model of applicant competition with endogenous effort is utilized to show that, in comparison to color-conscious affirmative action, these color-blind alternatives can significantly lower the efficiency of the student selection process in equilibrium. We examine data on matriculates at several selective colleges and universities to estimate the magnitudes involved. It is shown that the short-run efficiency losses of implementing color-blind affirmative action (in our sample) are four to five times as high as color-conscious affirmative action.

*We are grateful to Lawrence Katz, Kevin Lang, Steven Levitt, and Debraj Ray for helpful comments and suggestions. We also thank seminar participants at Boston University, Cornell, MacArthur's Social Interactions and Inequality Network, and New York University. Fryer is at the Harvard Society of Fellows and NBER, Littauer Center, Harvard University, Cambridge MA 02138 (e-mail: rfryer@fas.harvard.edu); Loury is at the Department of Economics, Boston University, 270 Bay State Road, Boston MA 02215 (e-mail gloury@bu.edu); and Yuret is at the Department of Economics, Boston University (e-mail: yuret@bu.edu). Financial support is gratefully acknowledged from the Andrew W. Mellon Foundation and the Carnegie Corporation.

"Implementing race-neutral programs will help educational institutions minimize litigation risks they currently face... If we are persistent in implementing race-neutral approaches, the end result will be to fulfill the great words of Dr. Martin Luther King Jr., who dreamed of the day that all children will be judged by the content of their character and not the color of their skin."

— U.S. Department of Education: *Race-Neutral Alternatives in Postsecondary Education: Innovative Approaches to Diversity*, Washington, DC., March 2003, pp. 7, 40

1 Introduction

The legal and political climate has shifted dramatically over the last decade on the issue of racial affirmative action. Accordingly, a number of institutions have begun to reformulate their policies – particularly in higher education. The states of Texas and Florida now guarantee admission to their public university systems for all in-state high school students graduating in the top ten and twenty percent, respectively, of their senior classes.[1] In the wake of Proposition 209 – a 1996 ballot initiative that banned racial affirmative action in California – public higher education officials there have substantially revised admissions practices.[2] Some private institutions have even decided to no

[1] In 1996, the state of Texas was ordered by a federal court to eliminate all race-conscious affirmative action in university admissions decisions. [See *Hopwood v. Texas*, 78 F.3d 932 (5th Cir. 1996).] The Texas legislature responded to *Hopwood* by passing House Bill 588, which guarantees Texas public high school students who graduate in the top 10% of their class admission to any Texas public college or university.

In February 2000, at the request of Governor Bush, the Florida State Board of Education banned consideration of race in admissions decisions for the state's higher education institutions. Florida's percentage plan, the Talented 20 program, took effect in August 2000. Under this plan, students who graduate from Florida's public high schools in the top 20% of their class, complete nineteen specific academic credits, and take an SAT or ACT test are guaranteed admission to one of eleven state universities, although not necessarily admission to the institution of the student's choice.

[2] The so-called "Eligibility in the Local Context" policy was implemented in California in the Fall of 2001. This program guarantees that the top 4% of each high school graduating class in the state will be admitted to one campus in the university system. For students admitted in the Fall of 2002, the UC system implemented a "comprehensive review" policy, which permits each campus to set admissions standards based on ten academic and four non-academic supplemental criteria, two of which may relate to socioeconomic status.

longer require that applicants submit standardized test scores.[3] A number of scholars and policy analysts have urged elite colleges and universities to rely more on the socioeconomic background and other non-racial, non-academic characteristics of prospective students when assessing their applications.[4]

Many justifications can be offered for these changes in admissions practice, but a primary factor would seem to be the desire to enhance racial diversity amongst the admitted without recourse to the use of explicit racial preferences. For this reason, we call these types of policies "color-blind affirmative action," in contrast to the more conventional, "color-sighted"affirmative action policies. This paper develops a simple model and explores some policy simulation exercises in order to study the economic implications of this emergent practice. We are particularly concerned with the question of whether, and the extent to which, a widespread shift toward color-blind affirmative action might be expected to impair the efficiency of resource allocation in higher education. The answer to this question is of considerable importance for public policy.[5]

It is important to understand that when one insists on color-blindness one does not thereby preclude the introduction of policies that favor a particular racial group. This is the point being made by the US Department of Education in the quotation above (advocating that conventional affirmative action be abandoned in favor of so-called "race-neutral programs.") Under color-sighted affirmative action, selectors give an explicit preference to individual applicants from some targeted racial group. A commitment to color-blindness prohibits such behavior. Even so, group-preferential goals can still be pursued tacitly under color-blindness, via policies that are implementable in a group-neutral manner. *Specifically, by exploiting their knowledge of differences between the race-*

[3]Mount Holyoke College, for example, has abrogated that requirement, while committing itself to admit some of the applicants who do not submit scores.

[4]Richard Kahlenberg (1996) is perhaps the most prominent advocate of so-called class-based affirmative action policies.

[5]It will be obvious in what follows that the ideas studied in this paper are of quite general relevance. "Color-blind affirmative action" arises in many areas of public policy having nothing to do with enhancing racial diversity. For example, a powerful legislator may want to influence the formula specifying how some public benefit will be distributed among juridisctions, with an eye toward benefiting his own constituency without appearing to be doing so. More generally, category-blind preferential policies can be used to pursue many group-redistributive goals (among population segments defined in terms of age, religous belief, gender, health status, region, nationality, and so forth), when decision-makers wish to avoid the appearance of playing favorites. We elaborate on this point in the Conclusion.

conditioned distributions of non-racial traits in the applicant population, selectors can favor a targeted racial group by over-emphasizing those non-racial productivity correlates that are relatively more likely to be found among members of that group.[6]

We investigate this problem by developing an analytical apparatus in light of which all of the aforementioned color-blind practices – percentage plans, voluntary test score submission, increased relative weight on non-test score criteria, preferential admission based on socioeconomic status – as well as conventional affirmative action policies, may be viewed within a unified framework. The central idea conceives of an abstract admissions policy function for a college or university. Given the applicant pool, this function maps each applicant's "profile" into a probability of admitting that applicant. An applicant's profile is merely a list of that applicant's "score" along a number of dimensions, not all of which need be directly related to academic achievement. In general, an applicant's chance of admission can be made to depend upon a host of factors. Conventional academic variables – test performance, grades in high school, recommendation letters, interview results – can be supplemented with information about an applicant's social background, life experience, geographic region of origin, extra-curricular interests, and the like. We think of the specific variables used in an admissions policy function, and the weights given to them, as being chosen by the college or university in order to meet its admissions objectives. An admissions policy function is said to be "color-blind" (or, equivalently, "race-blind") if, other things being equal, the probability of admission that gets assigned to a profile does not depend upon an applicant's race.

This paper derives two new insights from this general approach: First, by allowing (in our simulation exercises) for colleges to make their admissions decisions based on a vector of academic and non-academic applicant traits, we can deduce how a ban on color-sighted affirmative action might lower selection efficiency in the short-run, and alter the relative weights given to various factors (grades vs. test scores vs. socioeconomic background, for instance) in the selection process. Secondly, by studying (in our theoretical model) the ways in which color-blind policies distort applicants' incentives to acquire traits valued by selectors in equilibrium, we call attention to an additional loss in efficiency that may occur over the longer-run under these policies.

[6] Obviously, introducing a purely random element to the selection process can also raise the yield from any group that is statistically underrepresented in the pool of admittees. This point is stressed by Chan and Eyster (2003). However, one important contribution of this paper is to show that the options available to selectors for engaging in color-blind affirmative action are much broader that the simple use of randomization in the selection process.

There are two distinct ways in which color-blind affirmative action is inherently inefficient. First, in the short run, when the distribution of traits in the applicant pool may be taken as given, all affirmative action policies yield lower expected performance among the selected than does Laissez-faire. This is due to the fact that, under Laissez-faire (i.e., in the absence of any affirmative action policy), every admitted applicant is anticipated to perform better than any rejected applicant, which by definition cannot be true under any form of affirmative action. But, color-blind affirmative action is particularly inefficient in the short run, in the sense that its performance is always dominated by the best color-sighted affirmative action policy calibrated to achieve the same group representation goal. This is so because the non-racial factors which best promote selection from a targeted group are necessarily different from the non-racial factors which best predict post-selection academic performance – otherwise, some form of affirmative action would not be needed in the first place.

Secondly, color-blind affirmative action is likely to be inefficient over the longer run as well, when one considers how the distribution of traits presented by applicants will shift in response to the incentives created by colleges' admissions policies. Color-blind policies work by biasing the weights placed on non-racial traits in the admissions policy function so as to exploit the fact that some traits are relatively more likely to be found among the members of a preferred racial group. So, color-blind policies necessarily create a situation where the relative importance of traits for enhancing an applicant's prospects of being admitted diverges from the relative significance of those traits for enhancing an applicant's post-admissions performance. We show below that this is never the case under optimal color-sighted policy. Thus, to the extent that color-blind preferential policies distort applicants' decisions to acquire performance-enhancing traits prior to entering the selection competition, additional inefficiencies will emerge. We explore this possibility in our theoretical analysis.

Much has been written on the pros and cons of affirmative action, especially in the labor market.[7] However, until quite recently there had been little attention given in either the theoretical or empirical literatures to the problem of resource allocation inefficiencies due to affirmative action in higher education.[8] Two recent contributions warrant to be mentioned. Chan and Eyster (2003)

[7] Coate and Loury (1993) develops a theoretical framework for analyzing the incentive effects of affirmative action in the labor market. Holzer and Neumark (2000) is a comprehensive and insightful review of the theoretical and empirical literatures on affirmative action.

[8] Datcher Loury and Garman (1993) is an exception. That paper argues empirically that racial preferences in

have independently made one of the observations which we stress here – namely, that a ban on affirmative action could induce colleges to use non-racially preferential means to pursue their racial diversity ends, with potentially detrimental consequences for the efficiency of student selection. They study a constrained-optimal admissions problem for a college that values both student quality and racial diversity, that can rank students based on a one-dimensional measure of student ability, but that is enjoined from using racial preferences. Similar to our theoretical findings presented below, Chan and Eyster conclude that the second-best optimal admissions policy generally involves randomization, with the net result being that a ban on color-sighted affirmative action could end-up lowering the average quality of the college's admitted class. However, they are unable to address long-run efficiency issues in a comprehensive manner because, unlike in the present study, their model treats applicant characteristics as exogenous and one dimensional.

In another recent paper, Epple, Romano and Sieg (2003) have also taken note of the fact that a prohibition on explicit affirmative action can be expected to alter a college's use of non-racial information in the admissions process. Their numerically simulated model complements ours by focusing on the supply side of the higher education market. They introduce a framework where colleges, which differ in their attractiveness to applicants, compete with one another for the most desirable students. They are thus able to address the important question (which we here ignore) of how the distribution of students across a quality-hierarchy of colleges would be affected by a ban on explicit racial preferences. However they also take the distribution of applicant traits to be exogenous, so their analysis focuses on a different set of issue than those explored below.

The next two sections of this paper illustrate the workings of color-blind affirmative action by analyzing hypothetical college admissions problems, using data from the College and Beyond Survey (described in Bowen and Bok, 1998.) First, we introduce a formal analytical framework, show how the optimal color-blind affirmative action policy can be represented as the solution to a simple linear programming problem, and derive formulas for the optimal admissions policies in this setting. Then we report the numerical results of our policy simulation exercises, and examine how color-blind affirmative action would alter the relative weights given to non-racial traits in a

college admissions may induce an inefficient assignment of minority students to institutions (differentiated by their degree of selectivity). However, the evidence on this question is mixed. Using different data, Kane (1998) finds no support for the hypothesis of a detrimental mismatch for minorty students due to (color-sighted) affirmative action in college admissions.

college's admissions formula. We find that, compared to an optimal color-sighted selection formula, the optimal color-blind policy with the same representation goal puts less weight on academic factors and more weight on social background factors. These policy exercises also allow us to estimate the magnitude of the aforementioned short-run efficiency losses associated with color-blind affirmative action. We find that color-blind policies are significantly less effective at screening for the better academic prospects than are color-sighted policies calibrated to achieve the same racial representation target.[9]

Following this discussion of our simulation exercises, we present a simple theoretical model intended to illustrate the detrimental incentive effects of color-blind policies over the longer run. In the model, prior to applying for admission but anticipating the policies that colleges will employ, students make a binary effort decision that affects the distribution of their academic qualification. We show that, as the colleges' representation target approaches population parity, in the unique equilibrium of the model under color-blind affirmative action the fraction of students who choose high effort approaches zero. In contrast, under color-sighted affirmative action a goal of population parity can be achieved in equilibrium without vitiating students' incentives to choose high effort. This stark result graphically illustrates the potential for significant efficiency losses over the longer run due to reliance on color-blind policies. This paper can therefore be viewed as a theoretical and empirical evaluation of the limits of "race-neutral approaches" such as those advocated by the U.S. Department of Education in the passage quoted above. Our principle conclusion is that relying solely on "race-neutral approaches" to achieve greater racial diversity in higher education runs the risk of serious, unintended negative consequences.[10]

[9] The estimates reported here constitute a lower bound on the relative inefficiency of color-blind affirmative action, since we are not able with the data at hand to assess quantitatively the efficiency losses that may arise due to the detrimental incentive effects of these policies.

[10] For a more extended, critical discussion of these race neutral approaches – in the context of a specific legal dispute over the constitutionality of racial affirmative action at public universities – see the amicus curiae brief filed with the U.S. Supreme Court in the case *Grutter v. Bollinger* involving the University of Michigan Law School, Loury et al. (2003).

We realize, of course, that efficiency is not the only concern when assessing the desirability or the legality of alternative affirmative action policies. However, under the current Supreme Court's standards of legal scrutiny, a racial preference can be permitted if it constitutes a "narrowly tailored" means of furthering a "compelling state interest." Thus, once the goal of enhanced racial diversity in college admissions is acknowledged to be a compelling

2 Implementing Color-Blind Affirmative Action

To fix ideas, consider a concrete example of how color-blind affirmative action might work: Suppose initially that a college wants to admit a certain fraction of its applicants while maximizing the expected performance of those admitted. Let expected performance be a linear function of standardized test scores and of high school grades.[11] It is clear, then, that this college should adopt the policy of admitting only those applicants whose expected performance exceeds some threshold, where this threshold has the property that the fraction of applicants exceeding it just equals the fraction the college desires to admit. In effect, this means that the "weight" the college gives to grades relative to test scores in its admissions policy function should equal the ratio of the respective partial correlations of these variables with post-admissions performance. Now, suppose the college believes that following this threshold policy would lead to "too few" members of some racial group being admitted. Imagine that the college wants to obtain a greater degree of racial diversity while continuing to be race-blind in its treatment of individual applicants. Finally, suppose the college knows that among its applicants the distributions of grades within racial groups are much more similar to each other than are the corresponding distributions of test scores. Then the representation of the racial group with relatively lower (higher) test scores could be enhanced by setting the weight given to grades relative to test scores in the admissions policy function above (below) the level warranted by the relative correlations of these variables with post-admissions performance. To introduce such a change in admissions policy – as the states of Texas, Florida and California have, in fact, now done – would be to engage in the practice of color-blind affirmative action.

one, efficiency considerations become relevant to the legal determination of whether a given policy has been narrowly tailored to advance that purpose. A grossly inefficient policy, relative to some feasible alternative that achieves the same racial representation goal, is not a narrowly tailored one. See Ayers (1996), and the related discussion in Loury et al. (2003).

[11] We focus in the present example on two variables likely to enter any college's admissions policy function without intending to imply that these are the only variables of interest.

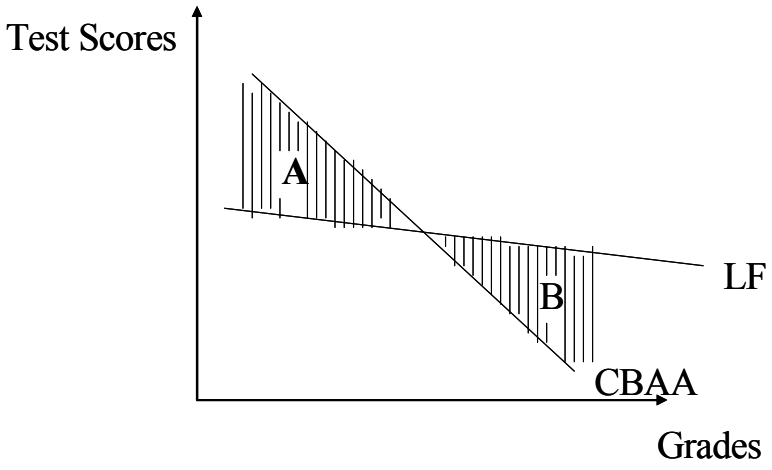

Figure 1: Relative Weight on Test Scores and Grades

Figure 1 captures the intuition at work here. The line segment LF represents a college's optimal admissions frontier under a policy of no affirmative action (call this Laissez-faire). Applicants above the line are admitted with probability one, while those falling below the line are admitted with probability zero. The line segment CB represents the same college's admission frontier under a policy of color-blind affirmative action. The CB frontier is steeper than the LF frontier because we have supposed that the distributions of grades within racial groups are more nearly equally to each other than are the distributions of test scores. The shaded area marked A in the figure depicts the set of applicants (with high test scores and low grades) who are rejected under CB, but who would have been admitted under LF. And, the area marked B shows the set of applicants (with high grades and low test scores) who are admitted under CB, but would have been rejected under Laissez-faire. Because the conditional probability that an applicant belongs to the targeted racial group is greater, given that the application falls in area B, than it is given that the application falls in area A, this college can enhance racial diversity in a race-blind manner by raising the weight it gives to grades relative to test scores when evaluating all applicants.[12]

[12] Note that this enhanced racial diversity is achieved at the cost of admitting a lower performing class on average, since the expected performance of every applicant in B is lower than that of any applicant in A.

Furthermore, suppose a college were to make reporting test scores optional for its applicants, while committing itself to admitting a certain fraction of its incoming class from the set of students electing not to submit scores (as Mount Holyoke College has, in fact, recently done.) In light of the incentives thereby created for applicants to

A. A Formal Framework

We now extend and formalize this example, by way of developing our simulated policy analysis in this paper. First we introduce some notation. Imagine that a college is to select an incoming class from a finite set of applicants. Let c denote the proportion of applicants to whom admission can be offered, $0 < c < 1$, and let r denote the target admissions rate for a disadvantaged minority group (relative to the size of the applicant pool). Let I be the set of all applicants, and take $i \in I$ to index a particular individual.

Suppose that each applicant belongs to one of two racial groups, and let $R_i \in \{1, 2\}$ denote the racial group membership of applicant i. Each application reports values for a bundle of non-racial traits (grades, social background factors, test scores and the like). Let J denote the set of non-racial traits, with specific traits indexed by $j \in J$. Then, the i^{th} student's application can be represented by the vector $(R_i; x_i)$, where $x_i \equiv (x_i^j)_{j \in J}$, and where x_i^j is the value which the application of student i reports for non-racial trait j. Moreover, the college's entire applicant pool can be represented by the (large) array $X = \{(R_i; x_i)_{i \in I}\}$.

Now, in general, an admission policy for the college in this setting associates with every applicant pool an array of probabilities specifying the chance that each applicant in the pool will be admitted. Let A_i be the probability of admitting applicant i, $0 \leq A_i \leq 1$. Then the college's admissions problem is to associate with each applicant pool, X, a vector of admission probabilities, $A(X) = (A_i)_{i \in I}$, so as to maximize the expected academic performance of the admitted class, subject to its capacity and racial representation constraints.

Let p_i be the college's expectation of the academic performance of applicant i. We assume that this expectation can be expressed as a linear function of the applicant's non-racial traits. Thus:

$$p_i \equiv [\text{Expected performance}|x_i] = \beta \cdot x_i = \sum_{j \in J} \beta_j x_i^j$$

for some vector of coefficients β. In addition, to the extent that the non-racial traits are distributed differently within the racial groups, a color-blind college could use this fact to estimate the racial group membership of any applicant presenting a particular vector of non-racial traits. Again,

selectively report their test scores, this too would be a color-blind policy that, although not explicitly preferential to any racial group, could be expected to result in more students from low-scoring groups being admitted.

adopting a linear specification of this relationship, we assume that:

$$r_i \equiv \Pr[R_i = 2|x_i] = \gamma \cdot x_i = \sum_{j \in J} \gamma_j x_i^j$$

for some vector of coefficients, γ.[13]

We will be considering the implications of three distinct policy regimes: Laissez-faire (LF); color-sighted affirmative action (CS); and color-blind affirmative action (CB). Consider the following simple linear program:

$$\max_{[A_i]_{i \in I}} \{(\frac{1}{c}) \sum_{i \in I} A_i p_i\}, \text{ subject to the following three constraints:}$$

(i) $A_i \in [0,1]\,,\, i \in I, \qquad$ (ii) $\frac{1}{|I|}\{\sum_{i \in I} A_i\} \leq c, \qquad$ (iii) $\frac{1}{|I|}\{\sum_{i \in I} A_i r_i\} \geq r.$

The maximand above is the anticipated average performance of the admitted class. Constraint (i) restrict the A_i to being probabilities; (ii) is the capacity constraint, and (iii) is the affirmative action representation constraint. In light of the discussion to this point, it should be clear that a college's optimal CB admissions policy must solve this linear programming problem. An optimal LF policy solves the same problem, but without constraint (iii). And, an optimal CS policy can be derived by first partitioning the applicant pool by race, and then solving parallel linear programs for each group, analogous to the LF version of the program above, but with the group-specific capacity constraints $r_2 = \frac{r}{1-\lambda}$ for group 2, and $r_1 = \frac{c-r}{\lambda}$ for group 1.

Solutions for the optimization problems implied by the three policy regimes are easily derived. Under the LF regime, one simply orders applicants by their expected performance, admitting the proportion c with the higher values of p_i. That is, for some number, μ (the Lagrangian multiplier on constraint (ii) above), we have that:

$$A_i^* = \begin{cases} 1 \text{ if } \beta \cdot x_i > \mu \\ 0 \text{ if } \beta \cdot x_i < \mu \end{cases}.$$

Here μ must be chosen in such a way that constraint (ii) holds with equality.[14]

[13] In what follows, we take it that the vectors of coefficients, β and γ, are known to the college and enter as parameters in its calculation of an optimal admissions policy. In our simulation exercises, we use data on matriculates at several institutions to estimate these coefficients, and then use these estimates to assess what optimal admissions policies might look like at these institutions.

[14] If $\beta \cdot x_i = \mu$, then A_i^* might need to lie strictly between zero and one for the capacity constraint to hold with equality.

Under the CS regime, there will be separate thresholds for the racial groups. So, for a pair of numbers μ_1 and μ_2, with $\mu_1 > \mu_2$, we have:

$$A_i^* = \begin{cases} 1 \text{ if } \beta \cdot x_i > \mu_{R_i} \\ 0 \text{ if } \beta \cdot x_i < \mu_{R_i} \end{cases} .$$

Here the μ_1 and μ_2 are to be chosen such that selection rates for the two groups are consistent with the capacity and representation constraints holding as equalities.

Under the CB regime, a Lagrangian multiplier on constraint (iii) alters the admissions policy relative to LF, because non-racial traits are now to be valued both for their association with prospective academic performance, and for their ability to predict an applicant's race. Thus, the optimal CB policy is characterized by two numbers, θ and μ', such that:

$$A_i^* = \begin{cases} 1 \text{ if } [\beta + \theta\gamma] \cdot x_i > \mu' \\ 0 \text{ if } [\beta + \theta\gamma] \cdot x_i < \mu' \end{cases} ,$$

where μ' and θ are such that both constraints (ii) and (iii) above hold as equalities.

This formalization captures nicely the ideas about color-blind policy mentioned in the introduction. Let j and k be two traits (high school grades and test scores, e.g.). Under LF and CS regimes, the college's "marginal rate of substitution" between traits j and k as reflected in the admissions policy function, denoted by $\text{MRS}_{j,k}$, is equal to the relative importance of these traits in forecasting student performance:

$$MRS_{j,k} = \frac{\beta_j}{\beta_k},$$

whereas, under the CB regime, the rate of substitution between traits j and k that holds constant the probability of being admitted is given by:

$$MRS_{j,k} = \frac{\beta_j + \theta\gamma_j}{\beta_k + \theta\gamma_k}.$$

These substitution rates are the signals sent out to applicants about the relative value of various traits in the admissions process. To the extent that the magnitude (and even the sign!) of such substitution rates is altered when color-blind means are used to pursue color-conscious ends, the incentives applicants have to acquire the relevant traits might be badly misaligned. In what follows we use data on student characteristics at selective public and private colleges and universities in the US to examine how color-blind affirmative action might be expected to play out in practice.

3 Simulating the Impact of Color-Blind Policies

We turn now to an application of the foregoing analysis in the context of college admissions in the US, using actual student profiles from the matriculating classes (entering college in 1989) of seven selective institutions (three research universities and four liberal arts colleges, labeled "College A" through "College G" in what follows). We conduct hypothetical admissions experiments, supposing that the colleges in question would have had to admit only half as many students as were, in fact, admitted. Their imagined selection problem is to choose which half of the students to retain from among the actual matriculates. The affirmative action goal is to maintain the original proportion of black students in this reduced class.[15] We estimate the loss of efficiency in selection that results from the imposition of the requirement to be color-blind in the selection process, given this racial representation goal. We also look at the nature of the constrained-optimal color-blind policies that emerge.

In this exercise, we compare the performance of our three policy regimes: LF (Laissez-faire); CS (color-sighted); and CB (color-blind affirmative action.) The capacity constraint for the colleges in these simulations is $c = 0.5$. The affirmative action representation target under CS and CB varies from college to college, as in each case the admissions policy maker seeks to maintain the same percentage of blacks among the selected students as had obtained in the original class. We employ the framework just discussed, modelling the constrained policy choices in each regime as linear optimization problems: The policy maker chooses an admissions policy, given the distribution of applicant traits and subject to capacity and representation constraints, so as to maximize the anticipated average academic quality of the admitted class. Once solutions for these linear programs are in hand for each college, we can compare the performance of the best admissions policy under each of the three regimes, and take note of how the constrained optimal color-blind policy attains its goal through an artful choice of racial proxies.

B. The Data

[15] While we focus here (for illustrative purposes) on the representation of black students, the color-blind methods we are discussing could be applied to promote (or restrain!) the presence of any identifiable subgroup of a student population. Table 6 below shows the impact on admission rates for Hispanics and Asians of color-blind affirmative action policies designed to foster the presence of blacks.

The College and Beyond database contains student level administrative data on college performance as well as information on admissions and transcript records of 93,660 full-time students who entered thirty four colleges and universities in the fall of 1951, 1976, and 1989. For the purposes of this paper, we restrict our attention to students from seven institutions in 1989. Our selection criterion is based solely on the availability of relevant data. Section 6 (the data appendix) describes how we combined and recoded some of the College and Beyond variables we use in our analysis. We employed three academic variables: SAT math score, SAT verbal score and High School Rank; and three socioeconomic background variables: mothers education, fathers education and the median household income of each student's zip code.[16] We use a simple regression analysis to associate these variables with the expected class rank after four years of matriculation achieved by the students in the sample (whose grade histories were available from the administrative records of the participating institutions). We also use these variables in a linear probability specification to estimate the conditional likelihood that a given student is black, given his or her non-racial characteristics.[17]

C. Results

Tables 1 and 2 present summary statistics for our sample for all students and for black students, respectively, broken down by institution. Black students score over one standard deviation below white students on the math and verbal sections of the SAT and have (on average) lower percentile ranks in high school, which is consistent with previous research. Among the socioeconomic variables, black students live in lower income zip codes and have parents who are less likely to be college educated. Table 3 reports results from the college-specific regression equations, which used academic and socioeconomic background variables to predict a student's class rank after four years of matriculation. Interestingly, both a student's SAT verbal score and their high school rank are

[16] Unfortunately, at some colleges family income information was only available for those students who applied for financial aid. So, we have used the median household income in the zip code of residence at the time of application as a proxy for socioeconomic status.

[17] Note that, while the estimated parameters β and γ are intended to apply to a college's applicant pool, the data on which our estimates are based come from matriculates, not applicants, at the various colleges. As a consequence, the coefficient estimates presented here could be seriously affected by problems of selection bias. However, because our goal with these simulation exercises is to assess the comparative short-run efficiency of color-blind and color-sighted policies, the conclusions we reach in this regard should be of interest despite this econometric difficulty.

stronger predictors of college rank upon graduation than is his or her SAT math score. Parental education is also a strong predictor of college rank. After controlling for our three academic variables and parental education, the average income of a student's zip code is not statistically significant. Table 4 reports results from the auxiliary regressions that we imagine the colleges to have run if, when operating under a color-blindness constraint, they needed to use academic and social background variables to forecast the likelihood that a student is black.

Tables 5 through 7 report the results of greatest interest, regarding the relative inefficiency of race neutral alternatives (T.5), the implication of such policies for the representation of Hispanics and Asians as well as Blacks (T.6), and the way that optimal color-blind affirmative action alters the weight given to various factors in the optimal admissions formula – i.e., test scores, grades and socioeconomic background measures (T.7)[18].

We measure the performance of a policy in terms of the average of the class rank predicted for the admitted students. Table 5 reports our estimates of the performance of CB and CS affirmative action policies at each college, relative to performance under Laissez-faire (which is normalized to be 100.) In addition, for comparison purposes the first two columns in the Table show the relative performance of LF policies when colleges are artificially prohibited from using information on SAT scores and high school grades, respectively. (Comparing these two columns, we note that SAT scores appear to be more valuable than high school grades for predicting post-admissions college performance in this sample, since prohibiting the use of test scores information reduces the performance of the LF policies by much more than does forbidding the use of grades.) The third and fourth columns in Table 5 report the relative performance of CS and CB admissions policies. As the theory predicts, in every case color-blind policies perform less well than do color-sighted policies. Moreover, the loss of selection efficiency associated with the use of color-blindness (relative to color-sighted policies intended to achieve the same representation target) is generally comparable to, and often exceeds, the loss of efficiency under a LF regime that arises when colleges are forbidden from using information on students' grades or test scores.

Table 6 shows the consequence of our three policy regimes (where the focus is on blacks) for

[18]The LF and CS optimal admissions policies both use the same weights (those derived from the regression predicting college class rank), while the CB policy employs weights that are "biased" in order to exploit the fact that some variables are more closely correlated (positively or negatively) than are others with a student's being black.

the overall ethnic/racial composition of the admitted class. Table 7 reports our calculations of the weights on students' characteristics in the admission formula that are employed under optimal LF-CS and CB policies. In effect, colleges are assigning a score to each student, and admitting that half of the applicant pool with higher scores. The numbers in Table 7 are simply the coefficients used in a linear formula to derive a student's score from that student's academic and socioeconomic traits. (LF and CS regimes use a common formula.) It is clear from the illustrative empirical results reported in Table 7 that optimal CB admissions policy gives less weight to test scores, more weight to high school grades, and more weight to social background factors than does optimal policy under the LF-CS regimes. As such, it is not surprising that CB admissions policy targeted on blacks tends also to raise admissions rates for Hispanics while lowering them for whites and Asians, as Table 6 reveals.

Our analysis of the data from these selective colleges and universities dramatically illustrates our principle finding: If an admissions office attempts to devise a system for achieving a desired degree of racial diversity while avoiding the explicit consideration of race, it would be significantly less effective at selecting students of all races who are anticipated to be high academic performers. The short-run efficiency losses (relative to Laissez-faire) from using race-neutral policies to pursue race-conscious goals are four to five times greater than the losses incurred when the same goals are pursued via explicitly race-sighted policies.

4 A Model of Incentive Effects with Endogenous Traits

As noted above, there are good reasons to expect that, because they distort applicants' incentives to acquire traits valued by selectors, color-blind policies may undermine the efficiency of the selection process over the longer run as well. Unfortunately, with the data at hand it is not possible for us to investigate empirically the potential magnitude of these long-run effects. However, by formulating a rigorous theoretical model of the selection problem with endogenous traits, we can gain some insight into the main issues.

Accordingly, consider the following simple model of competition for scarce positions, which builds on the framework used in the previous sections. A continuum of applicants (students) of unit mass consists of two racial groups, $R \in \{1, 2\}$, where $\lambda \in (0, 1)$ is the proportion belonging to

group 1. We will think of group 1 as "advantaged" relative group 2. Applicants seek to be accepted by any one of a large but finite number, N, of identical firms (colleges). Each applicant is randomly assigned to a firm, and so each firm faces an applicant pool of measure $1/N$ that is the statistical replica of the overall population. Let $\omega > 0$ be the gross value to an applicant of being accepted. Each firm can accept at most the fraction $c \in (0, 1)$ of those who apply. Firms prefer to accept the better-qualified applicants, and take the distribution of characteristics in their applicant pools as given, independent of their acceptance policies.

Prior to being assigned, applicants make an *ex ante* binary effort decision $e \in \{0, 1\}$ that affects their qualifications *ex post*. The incentive effects of affirmative action will be reflected in this model by the way that alternative policies alter the distribution in the student population of this binary effort variable, and the resulting distribution of qualifications. We assume that low effort ($e = 0$) is costless, but high effort ($e = 1$) entails a cost, $k \geq 0$, for an applicant. The frequency distribution of effort cost differs between racial groups. Group 2 is "disadvantaged" in the sense that it has a uniformly less favorable cost distribution than group 1. Let $G_R(k)$ be the fraction of group R with effort cost less than or equal to k , and let $g_R(k)$ be the associated density function. Further, let $G(k) = \lambda G_1(k) + (1-\lambda)G_2(k)$ be the CDF of effort cost for the overall population, with $g(k)$ being the associated population density function. The assumption that group 2 is "disadvantaged" can then be expressed as follows:

Assumption 1 : $[g_1(c)/g_2(c)]$ is a monotonically decreasing function of c.

In this model, an applicant's qualification (as perceived and valued by firms) will be a stochastic function of effort.[19] Let t be a number representing an applicant's qualification, let $F_e(t)$ be the probability that effort e leads to a level of qualification less than or equal to t, and let $f_e(t)$ be the associated density function. High effort is assumed to increase an applicant's qualification in the

[19] We emphasize that no asymmetry of information between firms and applicants is being assumed here. Qualifications are perfectly and costlessly observable by firms. Our assumption is that, when an applicant chooses effort *ex ante*, the extent of qualification that results *ex post* is random at the individual level. Because there is a continuum of applicants, the distribution of qualifications in any population is a function of the fraction of applicants in that population who choose high effort. Firms care only about an applicant's *ex post* qualifications, and not about the effort taken *ex ante*.

following sense:

Assumption 2 : $[f_1(t)/f_0(t)]$ is a monotonically increasing function of t.

Finally, let π_R represent the fraction of applicants in group R who choose action $e = 1$ ($R = 1, 2$), with π being the fraction of all applicants who exert high effort. The variables π_R are endogenous, and will depend on incentives for applicants to take high effort created by the firms' acceptance policies. Obviously,

$$\pi = \lambda\pi_1 + (1 - \lambda)\pi_2.$$

In a population where the fraction π exert high effort, the CDF of the distribution of qualifications is denoted $F(\pi, t)$, where

$$F(\pi, t) = \pi F_1(t) + (1 - \pi)F_0(t).$$

Let $f(\pi, t)$ denote the associated density function.

A. Firm Behavior

An applicant is characterized here by the pair (R, t), which specifies his or her group membership and degree of qualification. So, a firm's *acceptance policy* must be some function $A(R, t)$, representing the probability that an applicant with characteristics (R, t) is accepted.

Definition 1 *A firm's policy is color-blind if* $A(1, t) = A(2, t)$ *for almost every* t.

As before, in this theoretical discussion we consider the behavior of firms under three possible policy regimes: Laissez-faire, color-sighted and color-blind affirmative action (LF, CS, and CB, respectively). Moreover, in each case we study a decentralized situation where firms take π_R, the proportion of high effort applicants in each racial group, as given when deciding upon an acceptance policy.

Under LF, firms are unconcerned with diversity so they ignore group identity information. Thus, given π, firms choose the function $\{A(t) : 0 \leq t \leq 1\}$, with $0 \leq A(t) \leq 1$, so as to:

$$\max\{\int tA(t)f(\pi, t)dt\}, \quad \text{subject to } \int A(t)f(\pi, t)dt = c.$$

Trivially, the solution to this problem is given by: $\{A^*(t) = 1,\ t \geq t^* \text{and } A^*(t) = 0,\ t < t^*\}$, where $F(\pi, t^*) = 1 - c$. The best LF policy is a color-blind threshold policy, where firms accept the fraction c of the applicant pool with the highest qualifications.

Let r_2^* be the acceptance rate for group 2 that obtains under this LF optimal policy. We model affirmative action (either the color-blind or the color-sighted variety) by positing that firms seek an acceptance rate for group 2 members, r_2, that exceeds the "natural"rate, r_2^*, but is no greater than that implied by population parity. Thus:

$$r_2^* < r_2 \leq c.$$

Given their beliefs about the fraction in each group of applicants who have chosen $e = 1$, we require firms to choose an acceptance policy under which they anticipate to accept group 2 applicants at the rate, r_2. The aggressiveness of the affirmative action policy pursued by firms is taken to be exogenous throughout this analysis. In light of the capacity constraint, and given this two-group set-up, a representation target for group 2, r_2, necessarily implies a target for group 1, r_1, where:

$$\lambda r_1 + (1 - \lambda) r_2 = c.$$

Consider now a firm's selection problem under a CS policy regime, with representation target r_2. Taking (π_1, π_2) as given, firms choose $\{A(R,t) : 0 \leq t \leq 1\}$, with $0 \leq A(R,t) \leq 1$, so as to:

$$\max\{\int tA(R,t)f(\pi_R,t)dt\}, \quad \text{subject to} \int A(R,t)f(\pi_R,t)dt = r_R, \ R \in \{1,2\}.$$

Here the problem is solved separately for each group. As before, it is clear that the solution to this problem involves $\{A^*(R,t) = 1, \ t \geq t_R^*$ and $A^*(R,t) = 0, \ t < t_R^*\}$, where $F(\pi_R, t_R^*) = 1 - r_R$, $R = 1, 2$. Thus, under color-sighted affirmative action firms follow distinct threshold policies for each group, accepting the fraction r_R of the group R applicant pool with the highest qualifications, $R = 1, 2$.

Lastly, consider the firms' behavior in the CB regime, with the representation target for group 2 given as $r_2 < c$. Firms again take (π_1, π_2) as given, but now must choose a color-blind acceptance policy function so as to maximize mean qualifications of those accepted, while anticipating to generate the desired representation of group 2 members. So, the problem for firms becomes choosing $A(t)$ so as to:

$$\max\{\int tA(t)f(\pi,t)dt\}, \quad \text{subject to} \int A(t)f(\pi_R,t)dt = r_R, R = 1, 2$$

An equivalent way of expressing the firm's problem under a CB regime is as follows:

$$\max\{\int tA(t)f(\pi,t)dt\}, \quad \text{subject to } \int[A(t)f(\pi,t)dt] = c \text{ and } \int[A(t)f(\pi,t)\xi(\pi_1,\pi_2,t)dt] = (1-\lambda)r_2,$$

where

$$\xi(\pi_1,\pi_2,t) = \frac{(1-\lambda)f(\pi_2,t)}{f(\pi,t)}.$$

Thus, $\xi(\pi_1,\pi_2,t)$ is the conditional probability that an applicant belongs to the group 2, given group specific high effort rates (π_1,π_2), and given that the applicant's level of qualification is t. This problem is a linear program in an infinite dimensional space, as both the objective and the constraints may be regarded as linear functionals of the infinite dimensional control variable, $\{A(t) : t \in [0,1]\}$. Such programs are studied extensively in Anderson and Nash (1987), though we will see momentarily that we need not solve this problem explicitly to characterize the *equilibrium* distribution of applicant qualifications under a CB policy regime in this model.

Let μ' be the multiplier on the capacity constraint and let θ be the multiplier on the representation constraint in the firm's CB optimization problem stated above. Then, by the infinite dimensional analogue of the Kuhn-Tucker Theorem, the solution to this problem is given by: $\{A^*(t) = 1, t + \theta\xi(\pi_1,\pi_2,t) \geq \mu'$; and $A^*(t) = 0, t + \theta\xi(\pi_1,\pi_2,t) < \mu'\}$ Thus, an optimal acceptance policy function has the so-called "bang-bang" property. That is, for almost every t, either $A^*(t) = 1$, or $A^*(t) = 0$. It is also clear that, since $r_2^* < r_2$, this optimal policy under color-blind affirmative action cannot be a threshold policy. Hence, we may conclude that there will be levels of qualification t and s, with $t < s$, such that $A^*(t) = 1$ and $A^*(s) = 0$. *That is, imposing the color-blind constraint on firms that remain intent on achieving more representation for the disadvantaged group than occurs under Laissez-faire must lead to a situation in which some applicants are accepted while others with higher qualifications are rejected, thereby undercutting applicants' incentives to exert high effort.* This is a general feature of color-blind affirmative action policies, and it is the basic reason that such policies must be inefficient over the longer run, relative to the color-sighted alternative.[20]

[20] This bang-bang property would not hold if, in the manner of Chan and Eyster (2003), we were to impose some kind of monotonicity constraint on firms [e.g., requiring that $A(t)$ be non-decreasing, out of the incentive compatibility concern that applicants not see any gain from under-reporting their qualifications.] Still, the basic point we are making here would remain valid, even were we to impose monotonicity. Under such a constraint, the firm's problem can be reformulated so that it becomes (the dual of) what Anderson and Nash (1987, section 4.4) call a continuous semi-infinite linear program. If we apply their Theorem 4.8 (page 76) to this reformulated problem, we

Observe that any color-blind acceptance policy function, $A(t)$, defines two numbers: $A_e = \{\int A(t)f_e(t)dt\}, e \in \{0,1\}$. A_e is the probability that an applicant who exerts effort e is accepted, and $\omega[A_1 - A_0]$ is the expected benefit to an applicant from choosing action $e = 1$ instead of $e = 0$. So, all acceptance policies that generate a common value for $[A_0 - A_1]$ must induce the same behavioral response from applicants. Now, given that a firm's representation target for group 2 is r_2, and given that its capacity implies $c = \lambda r_1 + (1 - \lambda)r_2$, it follows that the firm's constraints under a CB policy regime can be written as follows:

$$\pi_R A_1 + (1 - \pi_R)A_0 = r_R, \ R = 1, 2.$$

From these equations it follows that $(A_1 - A_0) = \frac{[c - r_2]}{[\lambda(\pi_1 - \pi_2)]}$. Thus, under color-blindness, the assumption that workers face a binary effort choice implies that any feasible acceptance policy by firms engenders the same incentives for applicant effort, regardless of whether that policy is optimal. This observation greatly simplifies the task of finding equilibrium under color-blind affirmative action in this model.

B. Applicant Behavior

Regardless of the regulatory regime, an applicant with exogenous characteristics (R, k) who anticipates firms to employ the acceptance policy $A(R, t)$ will behave in accordance with the effort supply function, $e^*(R, k)$, given as follows:

$$e^*(R, k) = 1 \quad \text{if } \omega \int [A(R, t)[f_1(t) - f_0(t)]dt] \geq k, \text{otherwise } e^*(R, k) = 0.$$

That is, applicants choose to exert high effort if and only if their effort costs are below some threshold. Therefore, anticipating the acceptance policy function, $A(R, t)$, the proportion of group

can conclude that with a monotonicity constraint the firm's optimal acceptance policy can be expressed as a step function with at most two points of discontinuity. This, in turn, implies that there will be levels of qualification t and s, with $t < s$, such that $A^*(t) > 0$ and $A^*(s) < 1$. That is, some applicants are accepted with a probability strictly greater than zero, while others with higher qualifications are accepted with a probability strictly less than one, again undercutting applicants' incentives to exert high effort.

But, the main point we wish to emphasize is that, once applicants' qualifications are allowed to be endogenous in the manner that we follow here, the imposition of such a monotonicity constraint on the firm's acceptance policy is irrelevant for determining the distribution of qualifications in equilibrium. This is because (as we show below,) given the capacity and representation constraints, all feasible color-blind affirmative action policies for firms generate the same (diminished) effort incentives for applicants.

R who end-up choosing the action $e = 1$ is given simply by:

$$\pi_R^* = G_R(\omega \int A(R,t)[f_1(t) - f_0(t)]dt), R = 1, 2$$

C. Equilibrium

An equilibrium in this model is an acceptance policy for firms, $A^*(R,t)$, and an effort supply function for applicants, $e^*(R,k)$, that are mutual best responses to one another. We are now in a position to describe equilibrium behavior by applicants and firms under the three regulatory regimes. We require a bit more notation. Given a qualification threshold t and an acceptance capacity c, such that $F_0^{-1}(1-c) \leq t \leq F_1^{-1}(1-c)$, define $\widehat{\pi}(t,c)$ as the solution to the equation: $F(\pi,t) = 1-c$. That is, $\widehat{\pi}(t,c)$ is the proportion of an applicant population facing the qualification threshold, t, that would need to choose high effort if the acceptance rate in that population to just equal c. Obviously,

$$\widehat{\pi}(t,c) = [F_0(t) - (1-c)]/[F_0(t) - F_1(t)]$$

Moreover, let $\Delta F(t) = [F_0(t) - F_1(t)]$, $\Delta G(k) = [G_1(k) - G_2(k)]$, and define the function: $\delta(r_2) = \omega(c - r_2)/\lambda$. Finally, let π_R^* be in fraction of group R who exert high effort equilibrium, and let t_R^* be the qualification threshold that applies to group R in equilibrium (relevant only when firms employ a threshold policy.) It is straightforward to verify the following proposition.

Proposition 1 *Equilibrium may be determined as follows:*

(a) LF equilibrium is given by group-specific high effort rates (π_1^*, π_2^*) *and an acceptance threshold* t^* *such that:*

$$\widehat{\pi}(t^*, c) = G(\omega \Delta F(t^*)), \text{ and } \pi_R^* = G_R(\omega \Delta F(t^*)), R = 1, 2,$$

where firms' acceptance policy is $A^*(R,t) = 1$, $t \geq t^*$, *and* $A^*(R,t) = 0$, $t < t^*$, *and applicants' effort supply function is* $e^*(R,k) = 1$, $k \leq \omega \Delta F(t^*)$, *and* $e^*(R,k) = 0$, $k > \omega \Delta F(t^*)$.

(b) CS equilibrium is given by the pairs of group-specific high effort rates and acceptance thresholds, (π_1^*, t_1^*) *and* (π_2^*, t_2^*), *where:*

$$\widehat{\pi}(t_R^*, r_R) = \pi_R^* = G_R(\omega \Delta F(t_R^*)), \ R = 1, 2.$$

Firms follow the CS threshold policy $A^*(R,t) = 1$, $t \geq t_R^*$, *and* $A^*(R,t) = 0$, $t < t_R^*$, *and applicants follow the effort supply function* $e^*(R,k) = 1$, $k \leq \omega \Delta F(t_R^*)$, *and* $e^*(R,k) = 0$, $k > \omega \Delta F(t_R^*)$, $R = 1, 2$.

(c) CB equilibrium entails firms choosing some feasible acceptance policy and applicants choosing: $e^*(R,k) = 1$ *if and only if* $k \leq k^*$, *else* $e^*(R,k) = 0$, *where* k^* *solves* $\Delta G(k^*) = \delta(r_2)/k^*$

Proof. The claims (a) and (b) are a transparent consequence of the definition of equilibrium. Concerning (c), as noted above, capacity and representation constraints require that applicants face the effort incentive:

$$\omega(A_1 - A_0) = \omega[c - r_2]/[\lambda(\pi_1 - \pi_2)]$$

while optimal behavior by applicants implies that:

$$\pi_1 - \pi_2 = \Delta G(\omega(A_1 - A_0))$$

Combining these two equations to eliminate $(A_1 - A_0)$, and identifying k^* in (c) above with the value in equilibrium of $\omega(A_1 - A_0)$ yields the stated result. ■

D. The Impact of Affirmative Action on Applicant Qualifications in Equilibrium

Let us now consider how the pursuit by firms of greater representation of the disadvantaged affects the equilibrium distribution of the qualifications presented by applicants in the two groups.

We adopt the following notation: given firms' acceptance capacity, c, let $r_2^*(c)$ denote the proportion of group 2 applicants who are accepted in Laissez-faire equilibrium. Thus,

$$r_2^*(c) = 1 - F(G_2[\Delta F(t^*)], t^*), \text{where } t^* \text{ solves } \widehat{\pi}(t^*, c) = G(\omega \Delta F(t^*)).$$

We consider the impact of "weak" (that is, r_2 "close to" $r_2^*(c)$) and "strong" (that is, r_2 "close to" c) affirmative action goals. We wish to distinguish two cases under which the marginally accepted applicant in the absence of any kind of affirmative action has either a low or a high level of qualifications. Specifically, let t' solve the equation: $f_1(t')/f_0(t') = 1$. If in LF equilibrium $t^* < t'$ we say that firms' equilibrium acceptance standards are "loose" (because the marginally accepted applicant has low qualifications), while if $t^* > t'$ we say firms' acceptance standards are "tight"(because the marginally accepted applicant has high qualifications). Examination of Figure 2 should suffice to establish the following, which we state without proof.

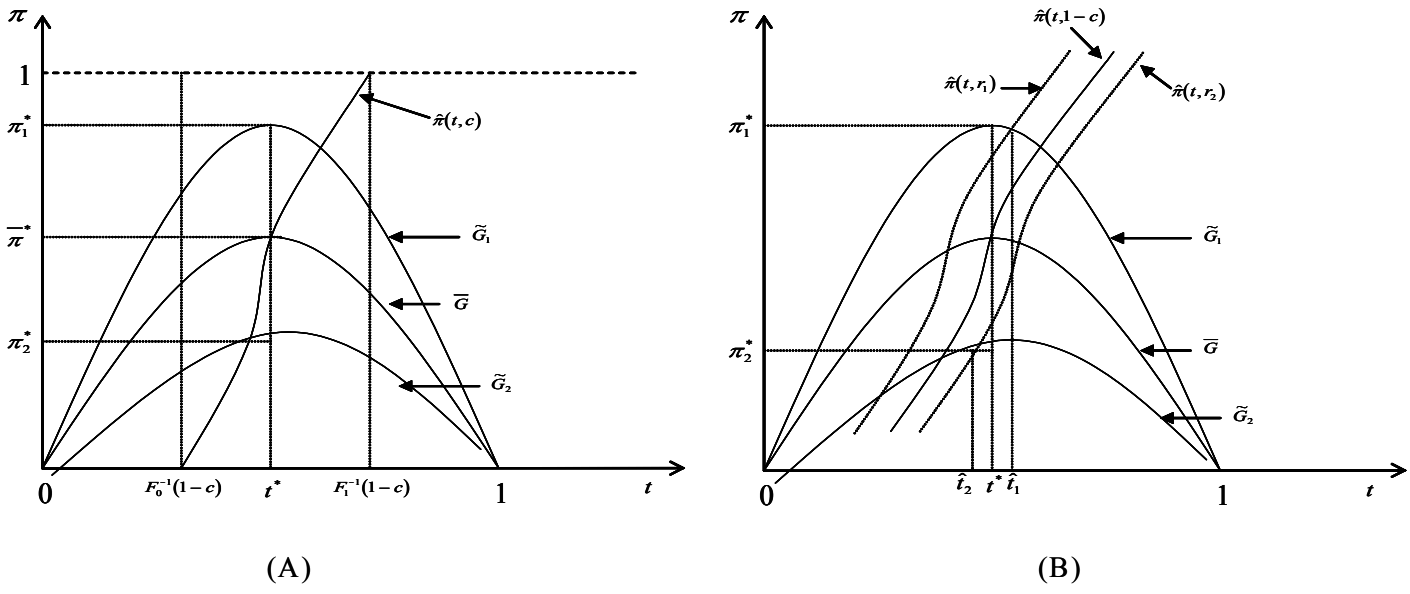

Figure 2: Laissez-faire (A) and Color-Blind (B) Equilibrium

Proposition 2 *If standards are loose in Laissez–faire equilibrium, then the pursuit of sufficiently weak CS affirmative action goals leads to an increase in qualifications among the advantaged group and a decrease in qualifications among the disadvantaged group, widening the racial qualifications gap relative to Laissez–faire. Moreover, tight standards in Laissez–faire equilibrium imply that weak*

CS affirmative action decreases qualifications among the advantaged, and increases qualifications among the disadvantaged, thereby narrowing the racial qualifications gap. When Laissez–faire equilibrium standards are close to the margin between being loose or tight then sufficiently strong CS affirmative action goals must decrease the qualifications of both groups.

Finally, consider the effect of color-blind affirmative action on applicant qualifications in equilibrium. To do so, we need one last definition. We will say that a selection problem is characterized by "elitism" if there is no acceptance policy by firms that can induce "high cost" applicants to exert high effort. (Here an applicant is said to have "high cost" if $k > k'$, where k' solves the equation: $g_1(k')/g_2(k') = 1$.) That is, elitism obtains when the structure of the situation is such that $\omega \Delta F(t') < k'$. (Note that effort incentives are maximal at the acceptance threshold $t = t'$, and that group disparity of qualifications is greatest when the marginal applicant to choose $e^* = 1$ has cost just equal to k'). Figure 3, together with Proposition 1(c), can be used to establish the following result:

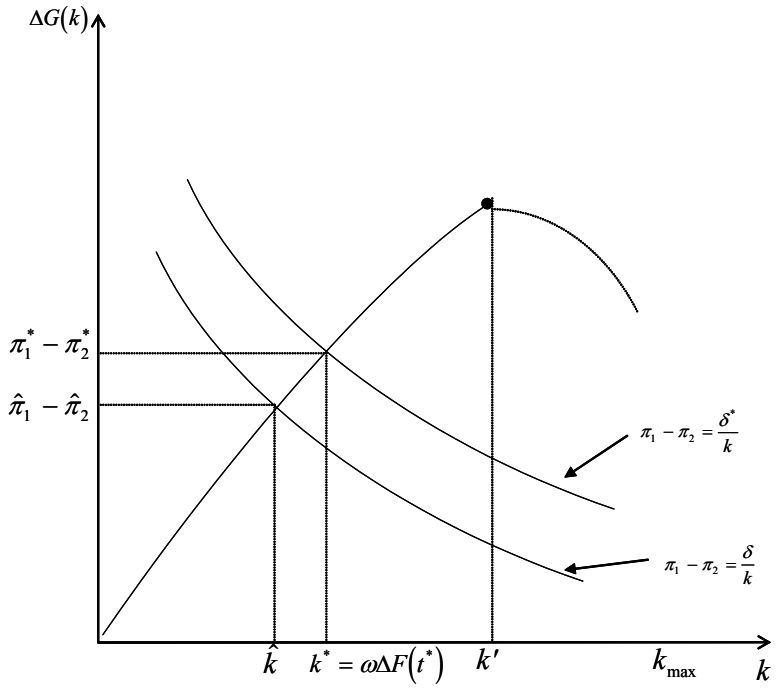

Figure 3: Color-Blind Affirmative Action Under Elitism, $\delta^* > \delta$

Proposition 3 *Suppose the condition of elitism obtains. Then for every affirmative action goal*

r_2, with $r_2^*(c) < r_2 < c$, there is a unique color-blind affirmative action equilibrium. Moreover, as r_2 rises, the level of qualifications in both groups declines, as does the qualifications gap between the groups. In the limit, as r_2 approaches c (population parity as a goal), the proportion of applicants choosing high effort in the color-blind affirmative action equilibrium approaches zero. However, there does not exist an equilibrium in this model under color-blind affirmative action that implements the representation target of population proportionality, $r_2 = c$.

This is a stark result that warrants further emphasis.[21] Consider an extreme example: one way to achieve population proportionality for all groups is to select from among candidates for a limited number of positions at random, with every applicant facing the same chance of success. This assures (with large numbers of applicants and statistical independence of applicant traits) that the fraction of successful candidates from any group equals the fraction of applicants from that group. Yet, random selection gives applicants no incentive to acquire traits valued by the selector. In equilibrium, the population of applicants (from all groups) will be much less distinguished under random selection, despite the fact that those selected will indeed be racially diverse. Proposition 3 demonstrates that the intuition of this extreme example extends to the general case.[22]

In light of Figure 3, all of the claims in Proposition 3 are straightforward, except the last on the impossibility of implementing an affirmative action goal of population parity in a color-blind fashion. Here is a demonstration of this result: We know that, as the r_2 approaches c, applicants' incentive to exert high effort vanishes. This leads to a non-existence result when $r_2 = c$ because, if all workers take effort $e = 0$, then the distributions of qualifications presented by workers in both groups in equilibrium would have to be identical. But then, no representation constraint would need to be imposed on firms to generate population parity. Yet, absent such a constraint, firms would want to behave as they do under Laissez Faire, accepting the fraction c of workers with the

[21] We have further results along these lines in a more general setting. See Fryer and Loury (2003), which considers the problem of equilibrium and optimal handicapping (i.e., affirmative action or, more generally, "categorical redistribution") in winner-take-all markets.

[22] There is an interesting externality here that promotes long run inefficiencies. Individual selectors drawing on a large, common pool of prospective applicants, may not take into account that their choice of selection criteria alters the distribution of traits in the overall applicant population. This makes the adoption of a random selection method look like a low cost move for any given selector. But when all selectors make this choice, they are all worse-off than they would be if none of them made it.

highest qualifications. Yet, were firms to do this, lower cost workers in both groups would then have an incentive to choose exert high effort, and the fact that group 2 is disadvantaged relative to group 1 would imply a failure to reach population parity in equilibrium. We conclude that when one group is disadvantaged relative to the other there can exist no color-blind equilibrium achieving population parity!

5 Moving Beyond Race

Our principle theoretical conclusion is that color-blind affirmative action entails a basic trade-off between incentives and representational goals. If firms are constrained to be color-blind but continue to value diversity, they will act in such a way as to "flatten" the function that relates a worker's probability of being accepted (in equilibrium) to that worker's level of qualification: Some lower qualification workers must have a greater chance of being accepted under color-blind affirmative action, and some higher qualification workers must have a smaller chance. (Otherwise, the disadvantaged group, which has relatively more low qualification members, cannot have its representation increased.) This flattening of the link between qualifications and success undercuts incentives for all workers to exert preparatory effort.

The issues explored in this paper are of more general interest, beyond the study of racial equity. There are many context in which a firm or public authority distributes some resource across a heterogeneous, categorically diverse population, with the dual objectives of allocating that resource to the most productive members of the population while avoiding an undue categorical disparity in receipt of the benefit.[23] A state government may need to distribute funds for public works among competing cities and towns, aiming to allocate the funds where they are most needed (or can best be made use of), while limiting any resulting disparity amongst jurisdictions. Similarly, a supplier of consumer credit (or insurance) may need to screen applicants according to creditworthiness (or insurability), without thereby generating a customer base with "too few" racial minorities. When the observable individual traits that are positively associated with creditworthiness (or insurability) are less frequently present in one population group than another, then simply screening-out the least qualified applicants could lead to a stark disparity in rates of selection between groups. For

[23] Akerlof (1978) investigates a related problem of "tagging" in the context of optimal taxation and welfare programs.

political, economic or legal reasons, such an outcome might be undesirable. However, it may also be undesirable in such settings to explicitly discriminate among applicants based on (race or sex-defined) group identities. This situation leads to the posing of an analytical problem nearly identical to the one investigated in this paper.

Alternatively, consider a customs union – e.g., the European Common Market. Imagine that a member state wants to favor its domestic producers of some good, but cannot do so directly without violating the trade agreement. Imagine further that members of the customs union are permitted to impose quality standards regulations which all goods, no matter where they originate, must meet. For instance, some Germans may want to limit imports into their country of Dutch beer, but may be forbidden to bar such products by Common Market rules. Still, they can require of any beer sold in Germany that it have so much hops, so little preservative, come in kegs that are made of a particular wood, etc.[24]

More generally, let the country have some preferences about what its quality standards should be, and suppose that the relative costs to domestic and foreign producers of meeting different standards are known. Suppose quality has two dimensions and that, compared to foreign producers, domestic firms are at an absolute cost disadvantage when forced to produce the Laissez-faire optimal quality vector. So, domestic producers would get a relatively low market share under the Laissez-faire optimal (i.e., disinterested) quality regulations. However, suppose domestic firms have a *comparative* cost advantage over foreigners in satisfying one dimension of quality. Then, by biasing regulation so as to give greater importance to that dimension of quality, the country in question can raise its domestic firms' market share without appearing to practice protectionism, but at the expense of having a less than optimal (given their natural preferences for quality) set of regulations. Again, we have arrived at a formulation analogous to the model studied here.

[24] Indeed, there is a real case involving Beck (German) and Heineken (Dutch) beers, in which the European Court of Justice prohibited Germany from enforcing its purity requirements for beer against beverages imported from other members of the EC. (See Commission of the European Communities v Federal Republic of Germany, Case 178/84, Judgment of 12 March 1987, 1987 ECR 1227.) We are grateful to William James Adams for bringing this example to our attention.

References

[1] Akerlof, George. 1978. "The Economics of 'Tagging' as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning," *American-Economic-Review*, 68(1): 8-19

[2] Anderson, E.J. and P. Nash. 1987. *Linear Programming in Infinite Dimensional Spaces: Theory and Applications*. New York, NY: John Wiley and Sons.

[3] Ayers, Ian. 1996. "Narrow Tailoring," *U.C.L.A. Law Review*, 43:6, 1781.

[4] Bowen, William and Bok, Derek. 1998. *The Shape of the River: Long Term Consequences of Considering Race in College and University Admissions*. Princeton, NJ: Princeton University Press.

[5] Chan, Jimmy and Eyster, Erik. 2003. "Does Banning Affirmative Action Lower College Student Quality?" *American Economic Review,* 93:2 (June), 858-872.

[6] Coate, Stephen. and Glenn C. Loury. 1993. "Will Affirmative Action Eliminate Negative Stereotypes?" *American Economic Review*, 83 (5), 1220-1240.

[7] Datcher Loury, Linda and David Garman. 1993. "Affirmative Action in Higher Education." *American Economic Review*, 83 (2), 99-103.

[8] Epple, Dennis, Richard Romano and Holger Sieg. 2003, "The Practice and Proscription of Affirmative Action in Higher Education: An Equilibrium Analysis," *NBER Working Paper* #9799, June 2003

[9] Fryer, Roland and Loury, Glenn. 2003. "Categorical Redistribution in Winner-Take-All Markets." mimeo. Harvard University.

[10] Holzer, Harry and David Neumark. 2000. "Assessing Affirmative Action," *Journal of Economic Literature*, 38:3 (September), 483-568

[11] Kahlenberg, Richard D. 1996. *The Remedy: Class, Race and Affirmative Action*. New York: Basic Books.

[12] Kane, Thomas J. 1998. "Racial & Ethnic Preferences in College Admissions," chp. 12 in C. Jencks and M. Phillips (eds.), *The Black-White Test Score Gap.* Washington D.C., Brookings Institution Press, 431-456.

[13] Loury, Glenn C., et al. 2003. "Brief of Social Scientists as Amici Curiae in Support of Respondents," in the Supreme Court of the United States, Barbara Grutter, Petitioner, vs. Lee Bollinger et al., Respondents, No.02-241 (www.bu.edu/irsd/MichiganBrief.doc).

6 Data Appendix

The College and Beyond data base contains admissions and transcript records of 93,660 full-time students who entered thirty four colleges and universities in the fall of 1951, 1976, and 1989. For the purposes of this paper, we restrict our attention to students from seven institutions in 1989. Our selection criterion is based solely on the availability of relevant data. All of our regressions include dummies for missing data. We describe below how we combined and recoded the variables used in our analysis.

1. SAT Math: Results of each student's performance on the mathematics portion of the Scholastic Aptitude Tests (SAT). The institutions reported each student's SAT scored. Only one SAT math score and one SAT verbal score was recorded for each student, even if the student took the test multiple times. Information is not available pertaining to which SAT score the institution reported.

2. SAT Verbal: Results of each student's performance on the verbal portion of the Scholastic Aptitude Tests.

3. High School Percentile: This captures the percentile rank of each student in their high school.

4. Parental Education: Parental education information was drawn from the student's college application. Questions involving parental education varied greatly from university to university. To account for this, we aggregated the data into two categories: college degree holder or not. This was done, independently, for mother's and father's education.

5. Zip Income: Student's residential Zip Code is drawn from College and Beyond dataset. We get median household income by zip code from 1990 Census.

Table 1: Summary Statistics For all Students

All Students	SAT Math	SAT Verbal	HS Percentile	Mother's Education	Father's Education	ZIP Income	# of students
College A	699 (66)	644 (73)	96 (6)	0.76 (0.43)	0.88 (0.33)	49780 (20480)	1139
College B	649 (80)	578 (76)	93 (10)	0.64 (0.48)	0.77 (0.42)	44149 (18317)	1858
College C	599 (86)	552 (84)	82 (16)	0.69 (0.46)	0.83 (0.38)	41256 (20839)	1486
College D	626 (91)	600 (94)	88 (11)	0.75 (0.43)	0.82 (0.38)	41772 (16940)	829
College E	595 (74)	580 (82)	90 (10)	0.73 (0.44)	0.84 (0.37)	42725 (16966)	606
College F	621 (81)	588 (82)	92 (9)	0.67 (0.47)	0.76 (0.43)	45518 (18617)	582
College G	682 (73)	655 (78)	94 (8)	0.71 (0.46)	0.79 (0.41)	46566 (19955)	519

Table 2: Summary Statistics for Black Students

Black Students	SAT Math	SAT Verbal	HS Percentile	Mother's Education	Father's Education	ZIP Income	# of black students
College A	612 (65)	595 (66)	95 (5)	0.63 (0.49)	0.66 (0.48)	39858 (20783)	82
College B	541 (82)	523 (83)	84 (15)	0.48 (0.50)	0.45 (0.50)	33223 (14568)	157
College C	515 (87)	502 (100)	84 (16)	0.68 (0.48)	0.75 (0.44)	26042 (11383)	98
College D	486 (94)	486 (105)	76 (18)	0.46 (0.50)	0.41 (0.50)	31792 (14492)	59
College E	515 (73)	523 (95)	84 (17)	0.69 (0.48)	0.65 (0.49)	34686 (15744)	24
College F	501 (74)	491 (72)	82 (16)	0.45 (0.50)	0.42 (0.50)	32308 (14401)	49
College G	563 (78)	552 (74)	84 (45)	0.48 (0.51)	0.55 (0.51)	37812 (18257)	33

Table 3: Predicted College Rank

College Rank	SAT Math	SAT Verbal	HS Percentile	Mother's Education	Father's Education	ZIP Income	Constant	R^2	N
College A	**7.67** (1.28)	**11.65** (1.15)	**7.64** (1.69)	**4.38** (2.07)	4.49 (2.73)	0.03 (0.41)	**-159.27** (16.89)	0.24	1139
College B	1.33 (0.88)	**6.08** (0.92)	**7.37** (0.75)	**3.71** (1.45)	**5.25** (1.69)	**1.65** (0.36)	**-75.04** (8.50)	0.13	1858
College C	**5.69** (0.87)	**6.58** (0.88)	**6.66** (0.53)	**6.64** (3.00)	2.50 (3.63)	**0.96** (0.34)	**-88.09** (7.33)	0.24	1486
College D	**4.05** (1.25)	**4.81** (1.22)	**4.26** (1.05)	2.59 (2.54)	**5.98** (2.86)	0.59 (0.60)	**-51.67** (10.24)	0.14	829
College E	0.45 (1.53)	**8.25** (1.36)	**8.39** (1.25)	**5.98** (2.97)	-0.25 (3.59)	-0.58 (0.77)	**-74.14** (15.08)	0.19	606
College F	**5.70** (1.57)	**7.09** (1.52)	**8.21** (1.73)	3.37 (2.99)	0.74 (3.34)	-0.11 (0.67)	**-103.85** (17.36)	0.17	582
College G	**8.87** (1.74)	**10.51** (1.57)	**8.82** (1.68)	4.25 (2.61)	4.62 (2.98)	0.81 (0.57)	**-173.41** (16.71)	0.34	519

Notes: College rank is percentiles in distribution of cumulative GPA among students who matriculated at that college in 1989. HS Percentile is students' percentile in his high school. Mother and Father's Education are dummies for students' mother and father being college educated. Zip Income is the average income of the student's zip code from the 1990 Census. Bold coefficients are five percent significant. Increments: SAT variables 100 points, HS Percentile 10 percentiles, ZIP Income 10000 dollars. We used dummies for the missing data. (Coefficients for these variables are not reported in this table).

Table 4: Probability of Being Black Regression

Black	SAT Math	SAT Verbal	HS Percentile	Mother's Education	Father's Education	ZIP Income	Constant	R^2	N
College A	**-12.31** (1.16)	-1.88 (1.00)	1.80 (1.53)	3.19 (1.88)	**-7.81** (2.48)	-0.36 (0.37)	**94.80** (15.50)	0.22	1139
College B	**-9.46** (0.81)	**-2.73** (0.84)	**-5.95** (0.68)	0.52 (1.32)	**-9.97** (1.54)	**-2.02** (0.33)	**156.40** (7.76)	0.22	1858
College C	**-6.42** (0.80)	**-3.59** (0.81)	**1.32** (0.49)	1.55 (2.79)	-1.41 (3.37)	**-1.87** (0.32)	**63.83** (6.81)	0.11	1486
College D	**-6.84** (1.04)	**-2.59** (1.02)	**-4.49** (0.87)	1.66 (2.11)	**-13.78** (2.38)	-0.93 (0.50)	**118.36** (8.51)	0.25	829
College E	**-4.69** (1.11)	**-2.36** (0.98)	-1.62 (0.90)	1.28 (2.15)	-3.60 (2.60)	-0.76 (0.56)	**65.38** (10.92)	0.07	606
College F	**-9.00** (1.35)	**-7.59** (1.30)	**-9.29** (1.48)	2.95 (2.55)	**-5.77** (2.85)	**-2.49** (0.58)	**207.56** (15.13)	0.34	582
College G	**-10.27** (1.58)	**-4.21** (1.43)	**-6.45** (1.53)	-0.83 (2.36)	-2.57 (2.70)	-0.35 (0.52)	**169.63** (15.17)	0.24	519

Notes: Dependent variable is student's probability of being black times 100. HS Percentile is students' percentile in his high school. Mother and Father's Education are dummies for students mother and father being College Educated. Zip Income is the average income of the student's zip code from the 1990 Census. Bold coefficients are five percent significant. Increments: SAT variables 100 points, HS Percentile 10 percentiles, ZIP Income 10000 dollars. We used dummies for the missing data. (Coefficients for these variables are not reported in this table).

Table 5: Relative Performance of Color-Blind and Color-Sighted Policies

	Relative Performances:			
	Laissez-Faire w/o SAT	Laissez-Faire w/o HS Percentile	Color-Sighted	Color-Blind
College A	92.23	99.50	98.85	94.23
College B	97.97	97.73	98.92	93.78
College C	96.40	95.80	99.85	98.59
College D	96.46	99.09	98.66	91.61
College E	97.04	96.98	99.83	97.72
College F	95.20	96.66	98.18	92.87
College G	93.78	99.50	97.80	85.30

Notes: Predicted College Rank of a student is estimated by the OLS regression. (Regression Results are given in Table 3). For each policy, we compute the average predicted College Rank of the admitted class. We call this value, the Performance of the Policy. To compute the Relative Performance, we index Laissez Faire's Performance as 100. For example color-sighted Relative Performance=(color-sighted Performance *100)/(Laissez-Faire Performance)

Table 6: Diversity

	Blacks			Hispanics			Asians			Whites		
	Laissez-Faire	Color-Sighted	Color-Blind	Laissez-Faire	Color-Sighted	Color-Blind	Laissez-Faire	Color-Sighted	Color-Blind	Laissez-Faire	Color-Sighted	Color-Blind
College A	5	41	42	7	7	20	83	80	68	474	441	439
College B	11	78	74	8	7	13	185	175	126	724	668	714
College C	28	49	53	20	17	23	39	38	41	651	635	617
College D	2	29	29	2	2	8	46	44	42.41	357	332	319.59
College E	6	12	12	2	2	4.33	31	30	30	255	251	249.67
College F	1	24	27.03	5	4	10	77	72	54	199	182	188.97
College G	0	16	19	1	1	9	30	27	18	228	215	212

Note: This table shows the effects of Laissez-Faire, Color-Sighted and Color-Blind policies on the composition of the admitted class. Note that Color-Blind policy is targeted for black students only. The number of students are not integers under Color-Blind policy. This is because, the college assigns an admission probability for two students strictly between zero and one. The college has a linear programming problem: to maximize a linear function with respect to some linear constraints. The college assigns a value to each student (with respect to his expected college rank and probability of being black) and admits the students with higher values. However, on the margin the college is indifferent between two students. (two students have equal values). That is why admitting a fraction of these students is consistent with the result being optimal.

Table 7: The Weight on Students' Characteristics in the Admission Formula for Laissez Faire and Color-Blind Policies.

	SAT Math		SAT Verbal		HS Percentile		Mother's Education		Father's Education		ZIP Income	
	LF	CB	LF	CB	LF	CB	LF	CB	LF	CB	LF	CB
College A	7.67	-5.37	11.65	9.66	7.64	9.55	4.38	7.76	4.49	-3.79	0.03	-0.35
College B	1.33	-4.95	6.08	4.27	7.37	3.42	3.71	4.06	5.25	-1.37	1.65	0.31
College C	5.69	1.72	6.58	4.36	6.66	7.48	6.64	7.60	2.50	1.63	0.96	-0.20
College D	4.05	-1.40	4.81	2.75	4.26	0.68	2.59	3.91	5.98	-5.00	0.59	-0.15
College E	0.45	-5.43	8.25	5.29	8.39	6.36	5.98	7.59	-0.25	-4.77	-0.58	-1.53
College F	5.70	-0.17	7.09	2.14	8.21	2.15	3.37	5.29	0.74	-3.02	-0.11	-1.73
College G	8.87	-5.09	10.51	4.79	8.82	0.05	4.25	3.12	4.62	1.13	0.81	0.33

Notes: Increments: SAT variables, 100 points; HS Percentiles, 10 percentiles; ZIP Income, 10000 dollars.