Abstract

Affirmative action in college admissions creates diversity by replacing marginal majority applicants with marginal minority applicants. Majority citizens’ preferences over the level of diversity as well as the means of achieving it depend on their competitive positions in admissions. In a simple model of college admissions, we show that higher-ability majority citizens prefer a more diverse student body and are more likely to favor race-conscious admissions rules than lower-ability citizens. Our model explains, first, the recent trend in American colleges and universities of replacing race-conscious admissions policies with race-neutral ones that de-emphasize standardized-test scores and, second, why majority voters support banning affirmative action even though it does not significantly improve their admissions prospects.

1 Introduction

The use of affirmative action in American college and university admissions has recently come under increased scrutiny. In 1995, the Regents of the University of California banned race-conscious admissions. In the 1996 case Hopwood v. Texas, a panel of the Fifth Circuit Court of Appeals forbade race-conscious admissions at all public universities in Louisiana, Mississippi, and Texas. In the same year, California voters approved Proposition 209, which prohibits public colleges and universities from using race in any admissions or financial-aid decision. Since then a series of court rulings, state legislations, and ballot referenda has banned race-conscious admissions in Washington, Georgia, and Florida. Presently, several judicial challenges are pending or under appeal, and the U.S. Supreme Court is widely expected to rule on the issue in near future.
Banning the explicit consideration of race and ethnicity in admissions does not mean that they no longer matter. In the wake the bans, affected universities have begun to alter their admissions policies. The new policies differ in their details but all share the common feature of de-emphasizing standardized-test scores.\(^3\) In Texas, the state legislature has passed a law requiring that public universities (e.g., UT Austin or Texas A&M) admit any candidate who graduated in the top ten percent of her high-school class, where rank is determined solely by high-school GPA. In California, the Berkeley campus of the University of California changed its measure of academic achievement to include many factors other than SAT I scores. Because minority candidates tend to score lower than majority candidates on many standardized tests, especially the SAT I, de-emphasizing test scores raises minority enrollment.\(^4\) Despite the fact that these policies are designed to favor minority candidates, they all enjoy widespread support among majority voters. Indeed, other states acting with the benefit of hindsight on California and Texas’s experiences have enacted similar measures. For instance, the Florida state legislature passed the “One Florida” plan whereby the top twenty percent of each high school class is granted admission to any public universities. Presently, the University of California is considering a proposal to replace SAT I by other subject tests in admissions.

Why do majority voters who oppose affirmative action favor other admissions policies ostensibly designed to increase minority enrollment? A ban on affirmative action that reduces minority enrollment of course does increase majority enrollment. However, the number of minority candidates for admission to elite colleges is so small that any improvement in majority admissions prospects too minor to be a likely explanation for the bans’ widespread popularity. Neither is moral objection to race-based policies likely to explain the bans, for if explicit race-based admissions is immoral, why isn’t changing admissions criteria to deliberately benefit one race over another?

In this paper, we provide an explanation for the recent bans based on the distributional effects of affirmative action among majority voters. Several recent studies show that a diverse student body enhances the learning experience of majority students. But affirmative action can create diversity only by displacing marginal majority candidates to make room for marginal minority candidates. Majority candidates may oppose affirmative action not because they oppose diversity \(\textit{per se}\) but because they are they are the majority candidates who are displaced by it. Depending on the likelihood of being replaced by minority candidates, majority candidates with the same innate

---

\(^3\)See Chan and Eyster (2001).

\(^4\)In the context of college admissions, the term “minority” usually refers to African Americans, American Indians, Chicanos, and Latinos. Asians are not included because they are overrepresented.
preferences over diversity may act as if they have different preferences over diversity as well as the
means to achieve it.

In section 2, we introduce a simple model of college admissions. In our model, a society is
composed of majority and minority citizens, where majority citizens are divided into three socio-
economic classes that indexed by “income,” which is low, middle, or high. Each citizen applies to
a single college that cannot accept all candidates. Each citizen has an ability, which is simply her
income. The minority group has an income lower than the average majority income and therefore
a lower average ability level. Citizens take an admissions test, where each citizen’s test score is
simply her ability level plus some noise term. An admissions rule assigns a threshold to each group,
and any citizen who scores above her group’s threshold is accepted. Rather than being fixed, the
noise term in the admissions test is a choice variable; increasing the noise level is tantamount
to de-emphasizing standardized tests. An admissions rule consists of an admissions test and a
threshold for each group. Under affirmative action, the minority group’s threshold is lower than
the majority group’s. A majority citizen’s utility is her probability of being accepted times the
value of being accepted, which increases in the diversity of the entering class and her income level.
Majority citizens know their own income, society’s income distribution, and the distribution of
ability conditional on income; they do not, however, know their own test score on any particular
admissions test.

In section 3, we show that majority citizens’ preferences over diversity and admissions test
correlate with income, even though their innate preferences are the same. Majority citizens trade
off their probability of being accepted and the diversity of the entering class. Since test scores and
income are positively correlated, under affirmative action the cost of diversity is borne by low- and
middle-income majority citizens more than by high-income ones; that is, low- and middle-income
majority citizens who would be accepted without affirmative action are more likely than high-
income majority citizens to be displaced by affirmative action. As a result, fixing the admissions
test, low- and middle-income majority citizens prefer less diversity than high-income citizens with
the same innate preferences over diversity. They also prefer to achieve diversity through noisy tests
— which minimize the chance that they are displaced — rather than through explicit affirmative
action.

The results are consistent with the fact that “insiders” of elite colleges — students, alumni,
faculty, and administrators — generally support affirmative action more than the public.

In section 4, we analyze the decision to ban affirmative action. American colleges and universi-
ties enjoy a high degree of autonomy, but through court rulings or legislation the public can impose certain restrictions on admissions policies — banning affirmative action or requiring that a public university admit the top ten percent of all seniors of each high school in the state. We model the political process as a two-stage game. In the first stage, majority citizens choose whether to ban affirmative action by means of majority-rule voting. In the second stage, the college, governed by insiders, choose an admission rule that complies with the public’s affirmative-action policy. For simplicity, we assume that college insiders are dominated by higher-income majority citizens and choose a rule that reflects their interests. When affirmative action is allowed, the college chooses the most accurate admissions test and uses affirmative action to achieve the preferred level of diversity of the insiders. When affirmative action is banned, the college may choose a noisy test to promote diversity. Majority citizens may ban affirmative action not only because they prefer less diversity than insiders, but also because by forcing the insiders to choose a more noisy test has distributional consequences within the majority groups: more low- and middle-income candidates are admitted. To highlight this point, we show that majority citizens may prefer banning affirmative action to using affirmative action and directly choosing the number of minority citizens admitted. Thus banning affirmative action actually may have little if anything to do with reducing minority enrollment. Indeed, public support for bans is almost always accompanied by suggestions for replacing affirmative action with class-based affirmative action.

This paper builds on earlier work analyzing the effects of banning affirmative action. Chan and Eyster (2001) show that a ban on affirmative action intended to raise student quality may backfire and lower it instead because an admissions office that cares about diversity responds to a ban by using a noisy admissions test that does not select the best candidates from either the majority or minority group. This is consistent with recent changes in admissions policies at public universities in Texas and California. In that paper, the admissions office’s preferences as well as the decision to ban affirmative action are exogenous, whereas in this paper they arise endogenously from distributional politics. More importantly, we show that noisy admissions tests may not be Pareto inefficient, and they may increase diversity by distributing the costs of diversity more evenly among majority citizens.
2 Admissions

2.1 Candidates

Society is composed of two ethnic groups: a majority group, \(W\), and a minority group, \(N\). The majority group is divided into three subgroups according to income, \(\lambda\). The size of the majority group is \(W\), and the size of group with income \(\lambda_i\) is \(W_i\); each member of group \(i\) has the same income \(\lambda_i\). Throughout, we assume that \(\lambda_1 > \lambda_2 > \lambda_3\), and we refer to citizens of these groups as high, middle, or low types, respectively. The minority group has size \(N\), and all minority citizens have the same income \(\lambda_N\). We assume that \(\lambda_1 > \lambda_2 > \lambda_3\); and we refer to citizens of these groups as high, middle, or low types, respectively. The minority group has size \(N\), and all minority citizens have the same income \(\lambda_N\).

The higher a student’s parents’ income and better their educations, the higher the student's high-school grades and test scores. (See, e.g., Vars and Bowen (1998).) We capture this regularity by equating ability with income and using \(\lambda\) to denote both. As a result of their lower income, minority citizens as a group have lower academic ability than majority citizens.

There is a single college to which every citizen applies and, if admitted, matriculates. College is free, and the benefit a majority citizen receives from attending it depends on her ability and the diversity of the student body. Let \(n\) denote the fraction of the college students who belong to the minority group. The value of college education to majority citizens is denoted by \(v(\lambda, n)\) with \(\frac{\partial v}{\partial n} > 0\) and \(\frac{\partial^2 v}{\partial n^2} \leq 0\). Since minority citizens’ preferences play no active role in our analysis, we simply assume that they receive some fixed positive benefit from attending college. The utility of any citizen who does not attend college is normalized to zero.

Many educators believe that an ethnically diverse student body contributes to the value of education. For example, Bowen and Bok (1998) report a survey of matriculants of elite colleges in which fifty-five percent of white matriculants consider that the “ability to work effectively and get along well with people of different races/culture” as a “very important” skill in life, and in the same survey sixty-three percent of white matriculants believe that their undergraduate experience was of considerable value in their development of this skill.

2.2 Admissions Rules

The college has a fixed capacity of \(C < W + N\), so only a fraction of all citizens can be admitted. Each citizen takes an admissions test, and each group has a cutoff such that only those citizens who score above their group’s cutoff are admitted. The test is a noisy predictor of academic ability with the score given by \(t = \lambda + x\varepsilon\), where \(\varepsilon\) is random term with zero mean and finite variance, whose density and cumulative-distribution functions are \(\phi\) and \(\Phi\), respectively. We assume that \(\phi\)
is everywhere positive, twice differentiable, symmetric, and satisfies the following assumption.

\textbf{Assumption 1} $\phi$ is log-concave.

Many common distributions, including the normal, have densities that are log-concave.\(^5\) In our model, Assumption 1 is equivalent to saying that test score $t$ and income $\lambda$ are affiliated, which, in turn, implies that for each $x$ the expected ability of citizens scoring $t$, $E[\lambda|t,x]$, increases in $t$.\(^6\)

\textbf{Lemma 1} $\phi$ is log-concave if and only if for any $y > z$, $\frac{\phi'(t-y)}{\phi'(t-z)}$ increases in $t$.

Assumption 1 is a natural property of any meaningful test. If a test consists of a number of independent questions, and higher-ability citizens answer each question correctly with higher probability than lower-ability citizens, then the test score satisfies Assumption 1.

A test $t = \lambda + x\varepsilon$ is a more accurate than another test $t' = \lambda + x'\varepsilon$ if $x < x'$. We assume that there is an exogenous lower bound on $x$ denoted by $\underline{x} > 0$; $\underline{x}$ is therefore an upper bound on the accuracy of the test. Since $t$ and $x$ are isomorphic, we refer to the test $t = \lambda + x\varepsilon$ as the test $x$ henceforth. An admissions policy consists of a test $x \in [\underline{x}, \infty)$ and cutoffs for the majority and the minority group, denoted $t_W$ and $t_N$, that satisfy the capacity constraint.

We can interpret $x$ as a characteristic of the test or as noise that an admissions office deliberately adds to existing measures of candidates’ qualifications. For example, achievement tests (e.g., SAT II or GCSE in the British system) are better measures of academic knowledge than aptitude tests (e.g., SAT I), and high-school grades may be more “random” than standardized-test scores because schools have different standards. More generally, college admissions decisions are based in part on non-academic factors such as perceptions of “character” or “leadership” and socioeconomic backgrounds; thus, $x$ can also be interpreted as the importance of these factors relative to other measures of academic achievements. Of course, $x$ may also be pure noise added to candidates’ test scores like at Berkeley’s Boalt Hall, where admissions are based not on candidates’ exact LSAT score but rather the elements of a partition of test scores.

When $t_W = t_N$, the admissions rule treats candidates from the two groups identically. When $t_N < t_W$, the admissions rule uses affirmative action. We assume that $t_W \geq t_N$: setting a higher standard for minority candidates is illegal. As $\underline{x} > 0$, some higher-ability citizens are rejected in

\(^5\)See Bagnoli and Bergstrom (1989).
\(^6\)The random variables $X$ and $Y$ are affiliated if whenever $x' > x$ and $y' > y$, their joint distribution $f$ satisfies $f(x,y)f(x',y') > f(x',y)f(x,y')$. 

6
favor of lower-ability ones, but for any finite $x$ higher-ability citizens have a higher probability of admission than lower-ability ones do.

Under the admissions rule $(x, t_N, t_W)$, a citizen with ethnicity $G$ and income $\lambda$ is admitted with probability $1 - \Phi \left( \frac{t_G - \lambda}{x} \right)$. The fraction of class belonging to the majority group is

$$1 - n = \frac{\sum_i \left( 1 - \Phi \left( \frac{t_W - \lambda_i}{x} \right) \right) W_i}{C}$$

and the fraction belonging to the minority group is

$$n = \frac{\left( 1 - \Phi \left( \frac{t_N - \lambda_N}{x} \right) \right) N}{C}.$$  

It is clear from Equations 1 and 2 that for any $x$, any $n \in [0, 1]$ is achievable by choosing the right $t_W$ and $t_N$. Henceforth, an affirmative-action admissions rule is represented by a duplet $(x, n) \in [x, \infty) \times [0, 1]$, and $t_W$ and $t_N$ are treated as functions of $x$ and $n$ defined by Equations 1 and 2. If affirmative action is banned, then $t_N = t_W$ and therefore an admissions rule is simply $x_b \in [x, \infty]$, where the subscript $b$ denotes the ban.

Since minority citizens as a group have lower ability than majority citizens, they will be underrepresented in the class (i.e. $n < N$) so long as $t_W = t_N$. There are two ways to admit more minority citizens. One is to use affirmative action — setting $t_W > t_N$ — and another is to make the test less accurate — setting $x > x$. Minority citizens are underrepresented because they have lower academic ability, and so putting less weight on ability raises their representation. When $x = \infty$, $n = \frac{N}{N+W}$. While both affirmative action and randomization can achieve any $n < \frac{N}{N+W}$, they do so by rejecting different majority citizens. As a result, different types of majority citizens have different preferences over admissions policies.

3 Preferences over Admissions Rules

Majority citizens from the three different income groups may have different preferences over admissions rules both because they have different intrinsic concerns for diversity and because they are in different competitive positions for admission. In order to separate these two effects, suppose that all majority citizens have the same income $\lambda$. In that case, under admissions rule $(x, n)$, a citizen receives an expected utility

$$U(x, n) = \frac{C (1 - n)}{W} v(\lambda, n).$$
Since all citizens have the same ability, they have the same chance of admissions, \( C(1-n) \), regardless of \( x \), and hence they are indifferent over \( x \). As \( \frac{\partial v}{\partial n} > 0 \) and \( \frac{\partial^2 v}{\partial n^2} < 0 \), preferences are concave in \( n \), and citizen \( i \)'s preferred \( n \) is determined by the first-order condition

\[-v(\lambda, n) + (1 - n) \frac{\partial v(\lambda, n)}{\partial n} = 0. \tag{3}\]

A marginal increase in \( n \) raises a majority citizen’s value of education by \( \frac{\partial v}{\partial n} \), which she gets with probability \( C \frac{1}{W} (1-n) \), and reduces her chance of admission by \( C \frac{1}{W} \), where the value of education is \( v \). Thus, the ratio \( \frac{\partial v}{\partial n} \) measures a majority citizen’s willingness to pay for diversity. If \( \frac{\partial v}{\partial n} \) increases in \( \lambda \), then the higher the majority citizens’ income, the higher their demand for diversity.

In order to isolate the effect of competition on preferences, we assume henceforth that majority citizens’ preference for diversity is income-neutral.

**Assumption 2** For each \( n \), \( \frac{\partial v(\lambda, n)}{\partial n} \) is independent of \( \lambda \).

Assumption 2 implies that high types do not prefer more diversity because of “wealth effects.” It holds, for example, when \( v(\lambda, n) = \lambda n^\alpha \), for some \( \alpha < 1 \).

When majority citizens have different incomes, they may prefer different levels of diversity even when Assumption 2 holds. Note that

\[
\frac{\partial U_i}{\partial n} = -\phi \left( \frac{tW - \lambda_i}{x} \right) C \frac{v(\lambda_i, n)}{\sum_j \phi \left( \frac{tW - \lambda_j}{x} \right) W_j} + \left( 1 - \Phi \left( \frac{tW - \lambda_i}{x} \right) \right) C (1-n) \frac{\partial v(\lambda_i, n)}{\partial n},
\]

where \( \frac{\phi \left( \frac{tW - \lambda_i}{x} \right)}{\sum_j \phi \left( \frac{tW - \lambda_j}{x} \right) W_j} \) is group \( i \)'s share of all majority citizens scoring at their cutoff \( t_W \), and \( \frac{(1-\Phi \left( \frac{tW - \lambda_i}{x} \right)) W_j}{\sum_j (1-\Phi \left( \frac{tW - \lambda_j}{x} \right)) W_j} \) is group \( i \)'s share of all admitted majority citizens. Given \( x \), admitting more minority citizens means increasing the cutoff \( t_W \) and therefore rejecting majority citizens scoring at \( t_W \). Hence, a citizen’s preference for diversity depends both on her probability of admission — the probability that she scores above \( t_W \) — and the probability that she scores at the cutoff \( t_W \).

Proposition 1 says that higher-income citizens have a stronger preference for diversity than lower-income ones in the sense that whenever lower-income citizens would like to admit more minority citizens, higher-income citizens do as well.

**Proposition 1** Given Assumptions 1 and 2, for each \( x \) and \( i < j \), if \( \frac{\partial U_i}{\partial n} \geq 0 \) then \( \frac{\partial U_i}{\partial n} > 0 \).

Intuitively, because higher-income citizens are less likely to score above rather than at \( t_W \) they attend more to the positive effect that diversity has on the value of education rather than the
negative effect it has on their chance of admission. Consequently, higher-income citizens favor more diversity. Formally, for any $\lambda > \lambda'$,

$$
rac{1 - \Phi\left(\frac{tW - \lambda}{x}\right)}{1 - \Phi\left(\frac{tW - \lambda'}{x}\right)} = \frac{\int_{tW}^{\infty} \phi\left(\frac{t - \lambda}{x}\right) dt}{\int_{tW}^{\infty} \phi\left(\frac{t - \lambda'}{x}\right) dt} = \frac{\int_{tW}^{\infty} \frac{\phi\left(\frac{t - \lambda}{x}\right)}{\phi\left(\frac{t - \lambda'}{x}\right)} dt}{\int_{tW}^{\infty} \phi\left(\frac{t - \lambda'}{x}\right) dt} < \frac{\phi\left(\frac{tW - \lambda}{x}\right)}{\phi\left(\frac{tW - \lambda'}{x}\right)}.
$$

The right-hand side of the second equation expresses $\frac{1 - \Phi\left(\frac{tW - \lambda}{x}\right)}{1 - \Phi\left(\frac{tW - \lambda'}{x}\right)}$ as a weighted average of $\frac{\phi\left(\frac{t - \lambda}{x}\right)}{\phi\left(\frac{t - \lambda'}{x}\right)}$ for $t \in [tW, \infty]$. As $\frac{\phi\left(\frac{t - \lambda}{x}\right)}{\phi\left(\frac{t - \lambda'}{x}\right)}$ strictly decreases in $t$, $\frac{\phi\left(\frac{tW - \lambda}{x}\right)}{1 - \Phi\left(\frac{tW - \lambda}{x}\right)} > \frac{\phi\left(\frac{tW - \lambda'}{x}\right)}{1 - \Phi\left(\frac{tW - \lambda'}{x}\right)}$. That log-concavity of $\phi$ implies increasing hazard rate.\(^7\)

Majority citizens’ income affects not only their preferences over diversity but also their preferences over the means to achieving any given level of diversity. Recall that minority citizens can be admitted through either affirmative action — setting $t_N < t_W$ — or by adding noise to the test — setting $x > x_0$. The following proposition says that for any given $n$, lower-income citizens prefer more randomization than higher-income citizens.

**Proposition 2** $\frac{\partial U}{\partial x} > (\geq) 0$ if and only if $E[\lambda|t_W] > (\geq) \lambda_i$. Furthermore, for each $i < j$ and $x < x'$, if $U_i(x', n) \geq U_i(x, n)$ then $U_j(x', n) \geq U_i(x, n)$.

To understand the first part of Proposition 2, it is instructive to compare a citizen’s probability of admissions under the admissions rules $(x, n)$ and $(x + dx, n)$ given the same $\varepsilon$. The more random test $(x + dx, n)$ raises the citizen’s test score and the cutoff by a small amount. If the citizen’s test score is significantly above the cutoff under the original test, then her new test score is above the new cutoff with the new test; likewise, if her test score is significantly below the cutoff under the original test, then her new score is below the new cutoff with the new test. Thus, the increase in noise only matters to those citizens scoring near the original cutoff $t_W(x, n)$. A citizen with ability $\lambda$ scores $t_W$ under $(x, n)$ only when $\varepsilon = \frac{tW - \lambda}{x}$. Under the new test $(x + dx, n)$ her score becomes $t_W + \frac{tW - \lambda}{x} dx$, and she is admitted and therefore better off if $t_W + \frac{tW - \lambda}{x} dx > t'_W$, where $t'_W$ is the new cutoff under $(x + dx, n)$.

\(^7\)This is equivalent to saying that $\Phi$ is log-concave when $\phi$ is log-concave. This fact was first proved by Prekopa (1972) and was introduced to the economics literature by Flinn and Heckman (1983). Bagnoli and Bergstrom (1989) presented a simple proof of the result. We include a proof similar to that of Bagnoli and Bergstrom (1989) to stress the connection between the increasing hazard rate of $\phi$ and the fact that $\frac{\phi\left(\frac{t}{x}\right)}{\phi\left(\frac{t}{x}\right)}$ increases in $t$ for any $y > z$. 9
To determine \( t'_W \), note that if the cutoff remained fixed, an additional \( W_i \phi \left( \frac{t_W - \lambda_i}{x} \right) \frac{t_W - \lambda_i}{x} dx \) citizens from income group \( i \) would be admitted by the new test. To satisfy the capacity constraint, the cutoff on the new test, therefore, must be raised by an amount equal to the average increase in test score among citizens scoring \( t_W \). Thus,

\[
\frac{\partial t_W}{\partial x} = t_W - E \left[ \frac{\lambda}{t_W} \right],
\]

where \( E \left[ \frac{\lambda}{t_W} \right] = \frac{\sum_{i=1}^{3} \phi \left( \frac{t_W - \lambda_i}{x} \right) W_i \lambda_i}{\sum_{i=1}^{3} \phi \left( \frac{t_W - \lambda_i}{x} \right) W_i} \) is the average ability of majority citizens scoring \( t_W \). Hence, a citizen with income \( i \) is better off under \( (x + dx, n) \) than under \( (x, n) \) when \( t_W + \frac{t_W - \lambda_i}{x} dx > t_W + \frac{t_W - E \left[ \frac{\lambda}{t_W} \right]}{x} dx \). Cancelling common terms yields the first part of Proposition 2. The second part follows immediately from the first. The intuition is straightforward: A slight increase in \( x \) affects only marginal citizens scoring \( t_W \). But conditional on the same test score, a lower-ability citizen must have a larger random term \( \varepsilon \) than a higher-ability one. As increasing \( x \) puts more weight on the random term, if the two citizens originally have the same test score, the lower-ability citizen will score higher than the higher-ability one in the new test.

Since for each \( x \) and \( n \), \( \lambda_1 > E \left[ \frac{\lambda}{t_W} \right] > \lambda_3 \), high types’ favorite test is \( x \), while low types’ is \( x = \infty \). Middle types prefer \( x = \infty \) when \( \lambda_2 \) is less than the mean, and \( x \) when \( \lambda_2 \) is higher than \( E \left[ \frac{\lambda}{t_W (x, n)} \right] \). For intermediate values of \( \lambda_2 \), middle types like best a test that is moderately accurate: accurate enough to reject low types, but not so accurate as to admit only high types.

**Corollary 1** Let \( x^*_i \) denote income group \( i \)’s preferred admissions rule. For each \( n \), \( x^*_1(n) = x \) and \( x^*_3(n) = \infty \), and

- \( x^*_2 = x \) if \( \lambda_2 \geq E \left[ \frac{\lambda}{t_W (x, n)} \right] \),
- \( x^*_2 \in (x, \infty) \) if \( \lambda_2 \in \left( E \left[ \frac{\lambda}{t_W (x, n)} \right], E \left[ \frac{\lambda}{x} \right] \right) \),
- \( x^*_2 = \infty \) if \( \lambda_2 < E \left[ \frac{\lambda}{x} \right] \).

Majority citizens do not have intrinsic preferences over the means to achieving diversity, and, by Assumption 2, their intrinsic preferences over diversity are independent of income. Yet, inter-group competition causes high types to behave as if they had a stronger preference for both diversity and affirmative action than low and middle types. A preference for diversity is closely related to one for affirmative action over randomization. High types like diversity more than low and middle types because conditional on scoring above \( t_W \), they are less likely to score at \( t_W \). For the same reason, affirmative action is their preferred means of achieving diversity: affirmative action displaces majority citizens scoring at the cutoff, while randomization rejects some majority citizens scoring above the cutoff as well. Our model thus predicts a positive correlation between preferences
for diversity and those for affirmative action. This distinguishes it from other explanations of systematic differences in attitudes toward diversity across social-economic classes. For example, if the high-income citizens like more diversity because they care more about “social justice”, then they should be indifferent between affirmative action and randomization.\footnote{The National Election Survey asks the following question “should the government in Washington see it it that black people get fair treatment in jobs or should the government in Washington leave these matters to the states and local communities”? People’s responses are highly correlated with their education level. Forty-three percent of those with at least a college degree strongly agreed, while only 31 percent of those without did.}

The following example illustrates how unequal access to college exacerbates the conflict over diversity between different social classes of the majority group.

\textbf{Example 1} $C = 0.5$. $\lambda_1 = 0.7$, $\lambda_2 = 0.2$, and $\lambda_3 = 0$. For all $i \in \{1, 2, 3\}$, $W_i = 1$. $\varepsilon \sim N(0, 1)$. $v(\lambda, n) = n^{0.2}$.

In Example 1, all three groups are of equal size, and the random term follows standard normal distribution. When all majority citizens have the same income, they prefer to allocate 16.7 percent of the class to minority citizens. Figure 1 depicts each type as a fraction of majority citizens admitted under various admissions tests when $n = 0.167$. Note that high types enjoy a huge advantage over the middle and lower types when the test is accurate, but the advantage diminishes rapidly as the test becomes noisy.

Figure 2 depicts the desired level of diversity of each type under various admissions tests. Note how unequal access to college education translates into divergent preferences for diversity. The horizontal line in Figure 2 represents the preferred level of diversity when all majority citizens have the same income. At the right end of the figure, different types have similar chances of gaining admission, and they prefer similar levels of diversity. Increasing the accuracy of the test improves
the competitive position of high types but worsens that of low types. As a result, high types prefer more diversity and low types prefer less as $x$ decreases. The desired level of diversity of middle types remains stable initially, but it declines rapidly as $x$ drops below 1. High types prefer more diversity when they are competing against middle and low types because it is mainly citizens of these other two types that affirmative action replaces. When the test becomes very accurate (i.e. when $x = 0.15$), few middle and low types are admitted. In that case, high types behave as if they are the only group in the society. In Figure 2, the level of diversity preferred by the upper class peaks at about $x = 0.55$.

4 The Political Process

American colleges and universities control their own admissions policies. Although the public cannot directly dictate admissions rules, it can impose certain restrictions such as banning affirmative action. We model the political process as a two-stage game. In the first stage, citizens from the majority group decide whether to ban affirmative action by way of majority voting. We exclude minority citizens from the voting process in the first stage solely in order to focus on majority citizens’ preferences. In the second stage, the college chooses an admissions rule to comply with the affirmative-action policy of the majority. The college is governed by insiders who graduated from the college in the past.\footnote{Imagine that citizens in the model are voting to maximize their children’s chance of admission.} As college admissions are based on test scores, insiders tends to have higher ability than an average majority citizens. As a result, they prefer more diversity and affirmative action than the public do. This is consistent with the fact many elite colleges support
affirmative action. For simplicity, we further assume that insiders share the same preference with the high types. It would not affect our qualitative results if the insiders’ preferences lie between high and middle types.

From Proposition 2, since for any \( t \) and \( n \), \( \lambda_1 > E[\lambda | t] \), high types always prefer as little noise as possible. If majority citizens decide not to ban affirmative action, the college sets \( x = \underline{x} \) and choose \( n \) to maximize

\[
U_1(x, n) = 1 - \Phi \left( \frac{t_W(x, n) - \lambda_1}{\underline{x}} \right) v(\lambda_1, n).
\]

Let \( n^*_h(x) \) denote the high types preferred \( n \) given \( x \).

When affirmative action is banned, the college can longer set \( x \) and \( n \) independently. For all \( x \) define \( \tilde{n}(x) \) such that

\[
t_N(x, \tilde{n}(x)) = t_W(x, \tilde{n}(x));
\]

\( \tilde{n}(x) \) is the fraction of class belonging to the minority group under a ban when the admissions test is \( x \). In this case, the only way that the admissions office can produce diversity is through \( x \). Recall that \( t_W(x, n) \) and \( t_N(x, n) \) are defined implicitly by Equations 1 and 2, respectively. Differentiating Equation 4 with respect to \( x \) yields

\[
\frac{d\tilde{n}}{dx} = \frac{\phi \left( \frac{t - \lambda_N}{x} \right) N}{x^2} \left( E[\lambda | t_{NW}] - \lambda_N \right),
\]

where \( t_{NW}(x) \equiv t_N(x, \tilde{n}(x)) \) is the common threshold under admissions test \( x \), and \( E[\lambda | t_{NW}] \) is the average ability of citizens, majority and minority, scoring \( t_{NW} \). The result is analogous to the first part of Proposition 2: an increase in \( x \) raises the representation of the minority group when their ability is less than the expected ability of a marginal citizen.

Assumption 3 \( C < \left[ 1 - \Phi \left( \frac{\lambda_1 - \lambda_N}{\underline{x}} \right) \right] N + \sum \left[ 1 - \Phi \left( \frac{\lambda_1 - \lambda_i}{\underline{x}} \right) \right] W_i. \)

The assumption means that there are insufficient seats to admit all citizens, including minority citizens, scoring above the mean test score for high types when \( x = \underline{x} \). It implies that \( t_W > \lambda_1 \).

Lemma 2 Given Assumptions 1 and 3, \( \frac{d\tilde{n}}{dx} > 0. \)

Since the minority group have lower ability than the majority group, increasing \( x \) on average increase minority enrollment. Lemma 3 says under Assumption 3 increasing \( x \) always average
increase minority enrollment. Assumption 3 ensures that the expected ability of a marginal citizen \( E[\lambda | t_{NW}(x)] \) is monotone decreasing in \( x \). Since \( \lim_{x \to \infty} E[\lambda | t_{NW}(x)] = E[\lambda] \), for all \( x \), \( E[\lambda | t_{NW}(x)] > E[\lambda] > \lambda_N \).

Under a ban, the college chooses \( x \) to maximize

\[
U_1(x, \tilde{n}(x)) = 1 - \Phi \left( \frac{t_{NW}(x) - \lambda_1}{x} \right) v(\lambda_1, n).
\]

Since high types’ ability exceeds the expected ability of a marginal candidate, raising \( x \) decreases their chance of admissions; this is the cost of randomization to the admissions office. The benefit of randomization is that it increases diversity and therefore the value of a college education. The college chooses \( x > \overline{x} \) under a ban if

\[
\frac{\partial U_1(x, \tilde{n}(x))}{\partial x} + \frac{\partial U_1(x, \tilde{n}(x))}{\partial n} \frac{dn(x)}{dx} > 0
\]

or

\[
\frac{dv(\lambda_1, \tilde{n}(x))}{v(\lambda_1, \tilde{n}(x))} \phi \left( \frac{t_{NW}(x) - \lambda_N}{x} \right) N \left( E[\lambda | t_{NW}(x)] - \lambda_N \right) = \frac{\phi \left( \frac{t_{NW}(x) - \lambda_1}{x} \right)}{1 - \Phi \left( \frac{t_{NW}(x) - \lambda_1}{x} \right)} (\lambda_1 - E[\lambda | t_{NW}(x)]).
\]

Intuitively, the college is more likely to choose \( x > \overline{x} \) when the income gap between the majority group and the minority group and that between the high type and the other two types are large. Let \( x_b \) denote the optimal test for the college under a ban. It is obvious that \( \tilde{n}(x_b) < n^*_1(x_b) \), the diversity level high types would have chosen given \( x_b \) when affirmative is allowed. But since \( n^*_1(x) \) may increase in \( x \) (as when \( x < 0.55 \) in Example 1), \( \tilde{n}(x_b) \) need not be less than \( n^*_1(x) \). It would however be the case when \( \overline{x} \) is large.

In the first stage, a majority citizen of type \( i \) votes to ban affirmative action if and only if

\[
U_i(x_b, \tilde{n}(x_b)) \geq U_i(x, n^*_1(x)).
\]

**Proposition 3**

Majority citizens vote to ban affirmative action if and only if \( U_2(x_b, \tilde{n}(x_b)) \geq U_2(x, n^*_1(x)) \).

High types always vote against a ban, since they get their preferred class with affirmative action. Proposition 3 means that whenever middle types prefer to ban affirmative action, so too do low types.

From Propositions 1 and 2, middle and low types prefer less diversity and more randomization than high types. Thus, they vote to ban affirmative action to force the college not only to lower diversity but also to increase \( x \). In the latter case, banning affirmative action may have nothing to
do with minority enrollment. To highlight this possibility, we show that majority citizens may ban affirmative action even when they can directly choose $n$.

**Proposition 4** Suppose that if affirmative action is not banned, then in stage 2 the college chooses $x$ and then majority citizens vote on $n$. Let $\lambda_3 < \lambda_2$ be given. Then there exists $K > 0$ such that for each $\lambda_1 > K$, affirmative action is banned.

Proposition 4 states that when high types score too much higher than middle and low types, middle and low types may prefer banning affirmative action to directly choosing $n$. Intuitively, when $\lambda_1$ is much higher than $\lambda_2$ and $\lambda_3$, few middle and low types will be admitted regardless of $n$. In this case, they are better off forcing the college to choose a larger $x$ by banning affirmative action.

Since only a small number of minority students gain admission to elite colleges under affirmative action, banning it cannot significantly improves the prospects of majority students. For example, at UC Berkeley the ban on affirmative action that took effect with the 1998 entering class reduced the number of admitted American Indian, African American, Chicano and Latino students by more than 50 percent. But the number of admitted minority students was small enough that this only increased majority candidates’ chance of admission by 3.5 percent; that is, if Berkeley had admitted just as many minority candidates in 1998 as it did when it used affirmative action in 1997, 26 rather than 29.5 percent of majority candidates would have been admitted. In subsequent years, Berkeley’s admissions attenuated the ban’s effect by changing its policies to admit more minority candidates: by 2000, a ban only increased average majority admissions prospects from 24 to 26 percent.

Kane (1998, page 453) suggests that majority voters may oppose to affirmative action because they overestimate the likelihood that an individual application may be affected by a ban.\footnote{Kane attributes the analogy to George Akerlof.}

Suppose that one parking space in front of a popular restaurant is reserved for disabled drivers. Many of the nondisabled drivers who pass by the space ... may be tempted to think that they would have an easier time finding a space if the space had not been reserved. Although eliminating the space would have only a minuscule effect on the average parking search for nondisabled drivers, the cumulative cost perceived by each passing driver is likely to exceed the true cost...

Proposition 4 explains why it may be rational for majority voters to ban affirmative action even when they as a group do not benefit much from it. That elite colleges in Texas and California react...
to bans by adopting race-neutral admissions rule that favor minority students may not discourage majority voters from seeking a ban in other states.

5 Appendix

Proof of Lemma 2: It is sufficient to show that $E[\lambda|t_{NW}(x)]$ decreases in $x$.

Write $E(\lambda|t_{NW}(x))$ as $\sum_{x} \omega_{i(x,n)}W_{i(x)}\lambda_{i}+$, where $\omega_{i}(x) = \phi \left( \frac{t_{NW}(x)-\lambda_{i}}{x} \right)$ and . Then,

$$\frac{\partial E(\lambda|t(x,n),x)}{\partial x} = \sum_{i=1}^{k} \frac{\partial E(\lambda|t(x,n),x)}{\partial \omega_{i}} \frac{\partial \omega_{i}}{\partial x}.$$ 

Note that $\frac{\partial E(\lambda|t(x,n),x)}{\partial \omega_{i}} = \frac{\lambda_{i}-E(\lambda|t,x)}{\sum_{i=1}^{k} \omega_{i}} > 0$ and $\frac{\partial \omega_{i}}{\partial x} < 0$ if $\lambda_{i} > E(\lambda|t,x)$. Thus every term is negative.

Proof of Proposition 3: We know that high types will never vote to ban affirmative action, so we need only show that when middle types prefer a ban, so too do low types. First note that $n_{b} < n_{1}^{*}(x)$: a ban lowers diversity. We need to show that if middle types prefer $(x_{b},n_{b})$ to $(x,n_{1}^{*}(x))$, then so too do low types. This happens when

$$\frac{\left(1 - \Phi \left( \frac{t_{b}-\lambda_{i}}{x_{b}} \right) \right)}{\left(1 - \Phi \left( \frac{t_{b}-\lambda_{i}}{x_{b}} \right) \right)} \geq \frac{v(\lambda_{2},n_{1}^{*}(x))}{v(\lambda_{2},n_{b})}.$$ 

Since

$$\frac{v(\lambda_{2},n_{1}^{*}(x))}{v(\lambda_{2},n_{b})} = \exp \left\{ \log \left( \frac{v(\lambda_{2},n_{1}^{*}(x))}{v(\lambda_{2},n_{b})} \right) \right\} = \exp \left\{ \int_{n_{b}}^{n_{1}^{*}(x)} \frac{dv(\lambda_{2},n)}{dn} \right\} = \exp \left\{ \int_{n_{b}}^{n_{1}^{*}(x)} \frac{dv(\lambda_{2},n)}{v(\lambda_{2},n)} \right\},$$ 

by Assumption 2 the right-hand side does not depend on $\lambda$, and we need only show that $\frac{1-\Phi \left( \frac{t_{b}-\lambda_{i}}{x_{b}} \right)}{1-\Phi \left( \frac{t_{b}-\lambda_{i}}{x_{b}} \right)} \geq \frac{v(\lambda_{2},n_{1}^{*}(x))}{v(\lambda_{2},n_{b})}$. Likewise,

$$\frac{1-\Phi \left( \frac{t_{b}-\lambda_{i}}{x_{b}} \right)}{1-\Phi \left( \frac{t_{b}-\lambda_{i}}{x_{b}} \right)} = \exp \left\{ \int_{x_{b}}^{\infty} \frac{-\phi \left( \frac{t(x)-\lambda_{i}}{x} \right)}{1-\Phi \left( \frac{t(x)-\lambda_{i}}{x} \right)} \left( \frac{\lambda_{i} - t(x)}{x^2} + \frac{1}{x} \right) dx \right\} = \exp \left\{ \int_{x_{b}}^{\infty} \frac{\phi \left( \frac{t(x)-\lambda_{i}}{x} \right)}{1-\Phi \left( \frac{t(x)-\lambda_{i}}{x} \right)} \frac{1}{x^2} \left( E[\lambda_{NW}|t] - \lambda_{i} \right) dx \right\}.$$ 

Since the hazard rate $\frac{\phi \left( \frac{t(x)-\lambda_{i}}{x} \right)}{1-\Phi \left( \frac{t(x)-\lambda_{i}}{x} \right)}$ is decreasing in $\lambda_{i}$, the integrand is decreasing in $\lambda_{i}$.
Proof of Proposition 4: As $\lambda_1 \to \infty$, $1 - \Phi\left(\frac{\lambda_2}{\lambda_1}\right) \to 0$ so $U_2(\lambda_2, n_2(x)) \to 0$ when affirmative action is not banned. When affirmative action is banned, high types choose $x > 0$, and so $U_2 > 0$. Thus majority citizens choose to ban affirmative action.

References


