Representation versus Assimilation: How do Preferences in College Admissions Affect Social Interactions?

Peter Arcidiacono* **Shakeeb Khan**† **Jacob L. Vigdor**‡

Duke University Duke University Duke University

May 25, 2010

Abstract

Given the existence of non-selective universities, the question of whether to employ racial preferences in college admissions reduces to one of optimal allocation of a finite resource: students who are members of under-represented racial or ethnic groups. In this paper, we assess recent legal arguments that racial preferences at selective colleges promote meaningful on-campus interracial interaction. As such, we model such interaction as a function of minority representation and, in some cases, perceived social similarity between students of different races. We estimate a structural model to capture these effects and use the results to trace out the net effects of racial preferences on population rates of interracial contact. The results suggest that the interaction-maximizing degree of racial preference, while positive, is significantly weaker than that observed in practice.

*Department of Economics, Box 90097, Durham, NC 27708. Phone: (919) 660-1816. Email: psarcidi@econ.duke.edu.

†Department of Economics, Box 90097, Durham, NC 27708.

‡Department of Public Policy Studies, Box 90245, Durham, NC, 27708. Phone: (919)613-7354. Email: jacob.vigdor@duke.edu. Gregory Besharov, William Darity, Kevin Lang, Tom Nechyba, Jesse Rothstein, Jeff Smith, Steven Stern, Sarah Turner, and Harriet Zuckerman as well as seminar participants at Brown University, Duke University, the Tinbergen Institute, Yale University, the NBER Higher Education Meetings, and the Winter Meetings of the American Economic Association provided helpful comments. We thank the Andrew W. Mellon Foundation for data access and the Smith Richardson Foundation for financial support. Any opinions expressed herein are those of the authors and not those of any affiliated institution. Teresa Romano provided excellent research assistance

Key Words: Affirmative Action, social stratification, diversity.

JEL: I2, J15, K0

1 Introduction

In 2003, two U.S. Supreme Court rulings (*Gratz v. Bollinger* and *Grutter v. Bollinger*) affirmed the constitutionality of certain types of racial preferences in college admissions.[1] The main premise behind the Supreme Court's decisions was that racial preferences in admissions benefit not only the minority students targeted by the policy, but majority students as well. One purported benefit to majority students is rooted in an increased likelihood of meaningful inter-racial contact associated with increased minority representation on campus.[2]

Given the existence of non-selective colleges, however, increasing the diversity of one campus necessarily entails reducing the diversity of another. Since only the most selective colleges employ racial admissions preferences, such preferences have little effect on the attendance rates of blacks (Arcidiacono, 2005). Bowen and Bok, in their seminal book *The Shape of the River*, write that:

> "Many people are unaware of how few colleges and universities have enough applicants to be able to pick and choose among them. There is no single, unambiguous way of identifying the number of such schools, but we estimate that only about 20 to 30 percent of all four-year colleges and universities are in this category. Nationally, the vast majority of undergraduate institutions accept all qualified candidates and thus do not award special status to any group of applicants, defined by race or on the basis of any other criterion." (Bowen and Bok 1998, pp. 16)

Thus, while racial admission preferences may increase diversity, and hence interracial interaction, at very selective colleges, there may be a countervailing loss of diversity and interaction at another campus. From a societal perspective, it is important to ask whether the net effect of this redistri-

[1] *Grutter v. Bollinger* let certain preferential policies stand whereas *Gratz v. Bollinger* declared certain ways of operationalizing racial preference to be unconstitutional.

[2] It has also been argued that higher minority representation at selective colleges allows more minority students to ascend to positions of power or influence in society, which may benefit majority students as well to the extent that the overall productivity of an organization depends on the degree of racial similarity between management and the more general workforce. Our paper is not capable of evaluating this claim.

bution of minority students is to increase or decrease the population level of interracial interaction. This paper seeks to address this question.

To this end, we first put forth some stylized facts about representation of blacks at selective schools, with the average SAT score of the school used as our measure of selectivity. Whereas a simple comparison of SAT score distributions by race would predict the lowest levels of diversity at the most selective colleges, we find that diversity is actually *higher* at these schools, relative to moderately selective universities.

Having presented these basic patterns, we introduce a simple model of collegiate behavior. The model both offers an explanation for the empirical regularities and illustrates the basic externality problem at the heart of the issue. When selective colleges act in their own self-interest, to raise the degree of interracial interaction on campus, they necessarily impose external costs on less-selective universities, where minority representation must decline. In some cases, these external negative effects more than fully offset the positive effects on the selective campus.

The net impact of affirmative action policies on the population rate of inter-racial contact depends critically on the process that governs whether such contact occurs. Is minority representation sufficient to guarantee inter-racial interaction, or do other personal characteristics matter? To consider this question, we first develop a series of models of inter-racial interaction, where interaction depends on the degree of minority representation on campus, and in some cases the degree of similarity between minority and majority students. The models make varying predictions regarding interaction rates and interaction maximizing admissions policies. None suggest that the interaction-maximizing relationship between percent black and college selectivity should be upward sloping. The models do suggest that in the quest to maximize contact between members of different groups, a tradeoff may arise between representation of the minority group and the degree of background similarity between members of different groups.

We test the empirical predictions of the models using data from the College & Beyond survey. The College & Beyond data set provides administrative records at thirty selective colleges in the United States as well as information from a follow-up survey conducted roughly seven years after matriculation.[3] This is the same data set used in Bowen and Bok's *The Shape of the River*. We analyze variation in self-reported measures of whether individuals who matriculated at a selective

[3]The sample restriction to selective colleges can be justified on the ground that the racial preference policies under examination here are only relevant at these schools (see Kane, 1998 and Arcidiacono, 2005).

college in 1989 became well acquainted with two or more members of various racial groups. We then relate these measures to the racial mix and characteristics of the school.

Empirically, we find substantial support for a model where similarity in academic background affects contact rates. We estimate models where the probability of inter-racial interaction is a function not only of the racial composition of one's cohort, but also of the degree of similarity between one's own test scores and those of other racial groups on the same campus. We find that individuals are most likely to interact with individuals of other races if those individuals have test scores similar to their own. In other words, students appear to socially stratify themselves by characteristics that correlate strongly with their test scores: cross-race interaction is most likely to occur within these strata. Increasing minority representation within the group of individuals having SAT scores significantly below one's own SAT score has very little impact on the probability of inter-racial interaction.

Evidence that likes interact with likes corroborates existing results in the emerging economics literature on friendship formation.[4] Foster (2005), Marmaros and Sacerdote (2006), and Mayer and Puller (2008) all find that race is extremely important in college friendship formation but find that similarity in academic background is also important.[5] Similarly, Weinberg (2005) analyzes high school students and also finds evidence for similarity in characteristics matters for friendship formation.

Reduced form results also suggest some scope for statistical discrimination. Increasing racial diversity of the group of individuals with SAT scores significantly higher than one's own SAT score has close to the same effect on inter-racial contact as increasing diversity of the group with scores similar to one's own score. This is consistent with a model where individuals prefer to interact with students of similar academic backgrounds and make inferences about peers' academic background based on their race. Well prepared black students may be penalized by this statistical discrimination as the presence of those with poor backgrounds has negative effects on the expectations of non-blacks.

[4]Likes interacting with likes has a much longer history in sociology under the heading of *homophily*. See Carli et al. (1991), Henderson and Furnham (1985), Kandel (1978), and McPherson and Smith-Lovin (1988).

[5]Camargo, Stinebrickner, and Stinebrickner (2009) suggest that part of the reason for race being so important is that students have misperceptions about their compatibility with those of other races. It is unclear how these misperceptions vary across college quality.

The reduced form results motivate a structural model that is designed to highlight the potentially offsetting effects of representation and similarity in academic background. The model has prospective friends of particular races arrive according to a process that is governed in part by the representation of those races at their school. Match quality is in part a function of the gap between one's own academic background and the background of the potential friend. Individuals receive noisy signals on the potential friend's academic backgrounds and use these signals as well as the school-specific distribution of academic background for members of the potential friend's race in assessing whether to form the friendship. Estimates from the structural model show that both representation and similarity in academic background are important factors in inter-racial friendship with some evidence of statistical discrimination.

With the structural estimates in hand, we perform a series of simulations to predict the amount of inter-racial interaction that would occur if the degree of racial preference in selective college admissions were adjusted to any of several points between currently observed levels and zero. Although race-blind assignments would reduce the representation of blacks at the most selective schools, the predicted population level of inter-racial interaction actually *increases*.[6] Decreases in contact at the most-selective colleges are more than offset by increases at moderately-selective colleges, owing in part to the reduction in academic background disparities on each campus.

Do these results imply that society would be best served by reducing the degree of racial preferences in selective college admissions? While the aim of this paper is not to answer this question, our work illuminates a trade-off that was not clearly articulated in arguments before Federal courts earlier in this decade.[7] If the intention of the U.S. Supreme Court was to promote interracial interaction only at the small set of colleges with the most stringent admissions standards, continuation of existing policy might make sense. If the social value of interracial contact is not strongly related to selectivity, however, current policies appear to be suboptimal.

The rest of the paper proceeds as follows. Section 2 provides a brief background on the dis-

[6] These simulations confine students to attend one of the institutions within the College & Beyond data set. If racial preferences were eliminated many of these students would not be admitted to any of these schools, reducing inter-racial contact. However, the students who did not attend one of these top schools would attend institutions further down the rankings and increase inter-racial contact at those schools. We discuss these effects in detail in section 7.

[7] It is also worthwhile noting that this paper is silent on the question of whether affirmative action is desirable from an equity perspective.

tribution of minority students across college campuses in the United States. Section 3 presents theoretical models of inter-racial interaction. Section 4 describes the data and methods used in the analysis. Section 5 examines the factors that influence inter-racial interaction. Section 6 describes and shows the estimates of the structural model. Section 7 shows the simulations of alternative assignment rules. Section 8 concludes.

2 Stylized Facts on Minority Representation in US Colleges

Given the existence of racial preferences in selective college admissions, one might expect the underlying relationship between minority representation and college selectivity to be strongly negative. After all, many professions or industries engage minority recruitment initiatives to address under-representation. Figure 1, which presents some basic information on the relationship between college quality and minority representation in the United States, shows that this intuition is at least partly wrong – over some range of the distribution, the relationship between selectivity and representation is positive. We use average SAT score as our measure of college selectivity, and percent black as our measure of minority representation. We use data from 1991 on 1,170 colleges garnered from the well-known US News and World Report rankings database.[8] The most selective among these colleges reported a mean SAT score of 1400; the least selective reported a mean of only 525. We regressed percent black of the school on a fractional polynomial of average SAT score using enrollment as weights.[9] Figure 1 plots this polynomial.

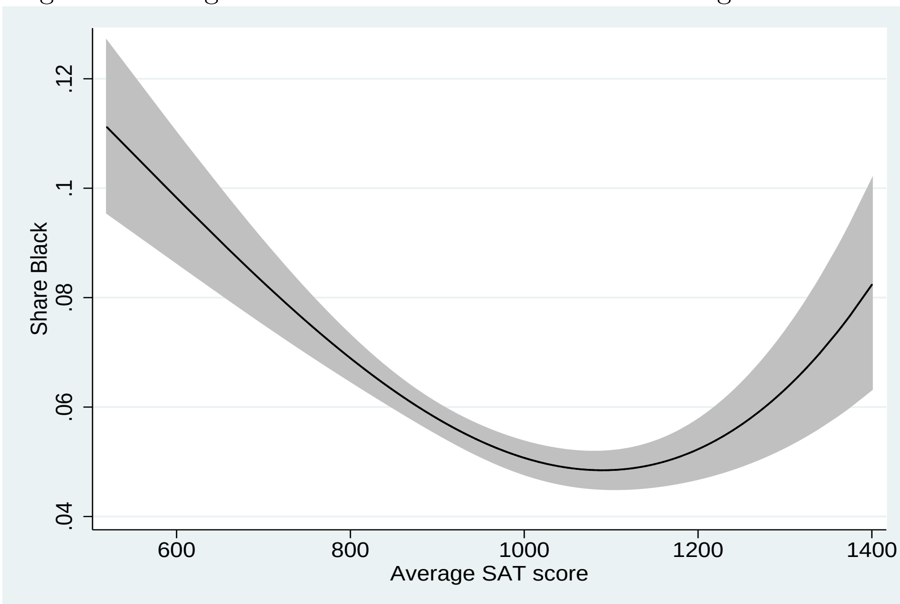
Plotting percent black as a function of average SAT score yields a U-shaped pattern: the most selective schools have a higher percent black than moderately selective schools. The minimum percent black is associated with an average SAT scores of 1090. Fourteen percent of schools (eighteen percent of the college student population) have average SAT scores greater than 1090. The mean rejection rate for schools with average SAT scores below 1090 is twenty-three percent and ninety percent of schools have rejection rates below 50%. Figure 2 displays the same table by region and in all regions the U-shaped pattern emerges, with the bottom of the U being lower in regions

[8] We are grateful to Dan Black for providing the data. U.S. News and World Report covers all colleges that are accredited and have more than 200 students. We exclude majority-black colleges from the analysis

[9] Fractional polynomials allow for both negative and non-integer powers, searching among different combinations of powers to best match the relationship between the two variables. Unweighted results showed the same pattern

with less diversity (the West and Midwest).[10] These U-shaped patterns hold despite blacks scoring significantly lower on the SAT than their non-black counterparts.

Figure 1: College Percent Black as a Function of Average SAT Score†



Source: America's Best Colleges produced by U.S. News & World Report. Weighted by college enrollment.

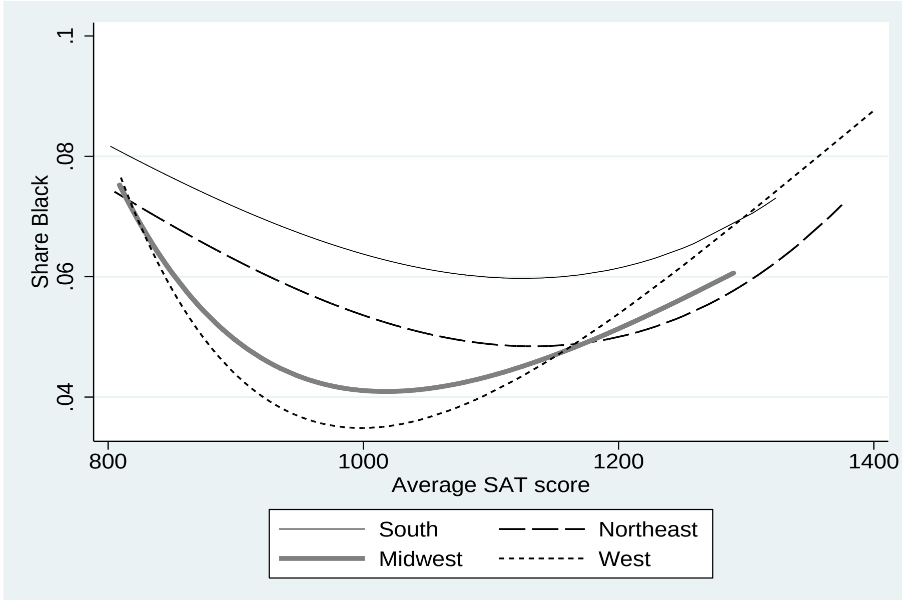
The U-shaped pattern suggests that preferences for blacks at the most selective institutions –schools in the top 10%– have come at the expense of diversity at the next-most-selective set of institutions. Indeed, racial preferences in admissions have been shown to have little effect on admissions at most schools (see Kane 1998 and Arcidiacono 2005) and in turn little effect on the overall attendance rates of blacks (Arcidiacono 2005). Race preferences in college admissions then primarily affect which colleges blacks attend rather than whether to attend at all.

3 A Model of Inter-racial Contact

Is the U-shaped pattern evident in Figures 1 and 2 consistent with some underlying model of optimizing behavior on the part of college administrators? Epple, Romano, and Sieg (2008) show

[10]We cut schools with SAT scores less than 800 out of the graph to make clear the differences in the minimum across regions.

Figure 2: College Percent Black as a Function of Average SAT Score by Region†



Source: America's Best Colleges 1991 produced by U.S. News & World Report. Weighted by college enrollment.

that the U-shape can result from competitive behavior by universities. Here we develop a simple model in the spirit of Epple, Romano and Sieg but also examine the implications of the model for inter-racial interaction.[11]

The model offers a formal treatment of the potential trade-off between raising the degree of inter-racial contact on selective university campuses and raising the population level of such contact. Intuitively, this is a basic example of an externality problem. When selective universities value inter-racial contact on their campuses, and when the minority group has lower average academic qualifications, they depart from a strategy of admitting only the most qualified students to accepting minority students with weaker credentials ahead of majority students with stronger credentials. Selective universities fail to consider, however, the effects that their decisions have on other campuses. The practice of affirmative action at elite campuses reduces the representation of

[11]The Epple, Romano, and Sieg model is significantly more complicated, including more universities, tuition, and university endowments. They do not examine interracial interaction within the university and how the interaction depends on the characteristics of the student body, the focus of our paper.

minority students on non-selective or less-selective campuses. The increase in inter-racial contact on selective campuses will therefore be offset, at least to some extent, by a decrease in such contact at less selective schools.

The model then shows conditions under which blacks will be a greater share of the student population at a more selective college than a less selective college. We presume that the more selective college is strictly preferred by all students, implying that this college's private optimization problem determines the equilibrium distribution of students across campuses. Under relatively straightforward conditions, we show that the global level of interracial contact increases when the more selective college relaxes its degree of preference for minority students.

Consider two universities, one of which is selective and the other is not.[12] Label these universities 1 and 2 respectively. There are two racial groups, B and W and within each group individuals either have academic background H or L. Denote π_R, π_{Rj}, and π_{Rj1} as the number of individuals of race R, the number of individuals of race R with academic background j, and the number of these who attend school 1, respectively. We then make the following assumptions about the size of the various groups and the size of the school:

1. The number of B types is less than the number of W types: $\pi_B < \pi_W$.

2. The number of B (W) types of academic background L is greater (less) than the number with academic background H: $\pi_{BL} > \pi_{BH}$, $\pi_{WL} < \pi_{WH}$.

3. The total population size is normalized to 2, $\pi_B + \pi_W = 2$, and the capacity of both schools is fixed at 1.

Assumption 1 establishes that B types are a minority, which will be a necessary condition for more B types to be at school 1 than at school 2 given the distributions of academic backgrounds (assumption 2). The assumption of stronger academic credentials within the W population is both consistent with available evidence and underlies the practice of affirmative action. Assumption 3 leads to school 1 and school 2 having the same number of students and also makes the math for the propositions below simpler.

[12] To get the full U-shape we would need an additional school. This makes the math more cumbersome without changing the ultimate conclusion that colleges acting privately fail to maximize the population rate of interracial contact.

Our next two assumptions assure the more selective university has the power to select its students and the option of enrolling only H types:

4. All individuals prefer to attend the selective school, school 1, but will attend school 2 if they are not admitted to school 1.

5. The number of H types is greater than one: $\pi_{BH} + \pi_{WH} > 1$.

Assumption 4 insures that school 1 gets to choose their student body before school 2. Assumption 5 guarantees that there are enough H types to fill the selective school.

We now specify how interracial interaction occurs. The next two assumptions imply that successful interracial interaction is more likely to occur between similar ability individuals.

6. At each school, individuals are randomly assigned one match, implying that the expected number of W types matching with B types at school 1 is $(\pi_{WL1} + \pi_{WH1})(\pi_{BL1} + \pi_{BH1})$ if the capacity constraint at school 1 binds.

7. The probability that a match is successful is 1 if individuals are of the same academic background and γ, $0 < \gamma < 1$, if they are of different academic backgrounds, implying that the number of successful inter-racial matches for WH types at school 1 is $\pi_{WH1}(\pi_{BH1} + \gamma\pi_{BL1})$.

Finally, we specify the objective functions of the the two universities:

8. The selective university's payoff function, U_1, and the non-selective university's payoff function U_2, depends linearly on the number of H types at their school and the number of successful inter-racial matches:

$$\begin{aligned} U_1 &= \theta_1(\pi_{WH1} + \pi_{BH1}) + \pi_{WH1}(\pi_{BH1} + \gamma\pi_{BL1}) + \pi_{WL1}(\gamma\pi_{BH1} + \pi_{BL1}) \\ U_2 &= \theta_2(\pi_{WH2} + \pi_{BH2}) + \pi_{WH2}(\pi_{BH2} + \gamma\pi_{BL2}) + \pi_{WL2}(\gamma\pi_{BH2} + \pi_{BL2}) \end{aligned}$$

where $\theta_1 > \max\{\theta_2, 0\}$ implies that H types are valued more at school 1 relative to school 2.

Altogether, the first seven assumptions are either innocuous or have a strong basis in reality. Assumption 8, which identifies the universities' objectives, is less obvious. We may infer from the behavior of selective universities in the United States, however, that enrolling more qualified students and a more diverse student body are both desirable (if sometimes conflicting) objectives.

The presumption that the true purpose of enrolling a diverse student body is to promote interracial interaction is a small leap, as there may be other rationales for enrolling a racially diverse student population. This objective is consistent, however, with objectives stated by universities in legal and policy debates over affirmative action. Finally, the assumption of a linear functional form is certainly debatable; the basic insight of this model holds, however, using more general formulations of the selective university's objective.

Given that B types are a minority (assumption 1) and that the school could fill up using only H types (assumption 5) it is clear that the selective university will choose to admit all of the B types of academic background H and none of the W types of academic background L: $\pi_{BH1} = \pi_{BH}$, $\pi_{WL2} = \pi_{WL}$. The selective university then solves the following maximization problem:

$$\max_{\pi_{WH1}, \pi_{BL1}} \quad \theta_1(\pi_{BH} + \pi_{WH1}) + \pi_{WH1}(\pi_{BH} + \gamma\pi_{BL1})$$
$$\text{s.t. } \pi_{BH} + \pi_{BL1} + \pi_{WH1} = 1$$

Proposition 1. *School 1 will have a higher fraction of B types than school 2 when the following condition holds:*

$$\frac{\pi_B}{2} < \frac{1}{2} - \frac{\theta_1}{2\gamma} - \frac{(1-\gamma)\pi_{BH}}{2\gamma}$$

Note that the second two terms are negative, which is why B types have to be a minority to be over-represented given the distribution of academic background. Note further that shifting the distribution of B types from L to H lowers the right hand side of the condition, making the condition *less* likely to hold. This occurs because adding B types with academic background L can crowd out more productive interaction between WH types and BH types.

While conditions exist for the more selective university have a higher fraction of B types than the non-selective university, inter-racial contact could be increased by moving trading some BL types at school 1 for some WH types at school 2:

Proposition 2. *Conditional on all BH types being assigned to school 1, the point where inter-racial contact is maximized will result in the fraction of B types at school 1 being less than the fraction of B types at school 2*

The intuition for the proposition is clear. If inter-racial interaction did not depend on similarity in characteristics, the interaction maximizing point would be to have the fraction of B types at selective school set equal to the fraction of B types at the non-selective school. When similarity in characteristics matter for interaction, then a slight shift from equal representation to more representation at the less selective school leads to more interaction due to the higher rates of success at the non-selective school.

Note that maximizing inter-racial contact ignores the fact that H types are better suited for the selective school.

Proposition 3. *A social planner maximizing the sum of the two universities' objective functions will reduce minority presence at school 1 relative to the inter-racial contact maximizing level in Proposition 2.*

Consider the case where the market operates and condition 1 holds. Relative to this outcome, a social planner interested in maximizing inter-racial contact will exchange some BL types at school 1 for some WH types at school 2 (Propositions 1 and 2). If a social planner also values matching H types to school 1, the social planner will exchange even more BL types at school 1 for WH types at school 2 (Proposition 3).

These results correspond to the classic negative externality scenario. Actors fail to consider the negative impact of their actions on other actors, and therefore engage in an excessive amount of a particular action. In this case, the selective university's preoccupation with levels of interaction on its own campus lead it to take actions that have substantial negative effects on the degree of interaction on non-selective campuses.

4 Data

To examine the impact of collegiate diversity on inter-racial interaction, we employ the College & Beyond Data set, made available by the Andrew W. Mellon Foundation.[13] This data set contains information from two sources: administrative information on a set of mostly selective undergraduate institutions, and survey responses collected from a sample of students who matriculated at

[13] We omit observations from historically black colleges as race preferences for under-represented minorities are not relevant at these schools. We also omit observations from women's colleges as the selection into and the environment of these colleges may be substantially different from other institutions. Including data from women's colleges does not affect our results.

those institutions in one of three cohorts. Our analysis focuses on the 1989 entering cohort.[14] The administrative data include information on each student's SAT scores, major, and their ultimate means of exit, whether graduation, transfer, or withdrawal. For most institutions, the administrative data represents the entire entering cohort. For the remainder, the data comprise a nonrandom sample of the student body. Weights are provided to adjust for this sampling. A complete list of institutions represented appears in the appendix. A subset of the administrative sample, approximately 40%, received and replied to a follow-up survey in 1996. This survey provides information on many outcomes of interest such as collegiate satisfaction and inter-racial interaction.

4.1 Evidence of admission preferences: disparities in academic background

As a proxy for academic background, we focus on SAT scores.[15] Table 1 shows the distribution of SAT scores for blacks and non-blacks relative to the school average. Relative to members of other racial groups, blacks in the College & Beyond received lower SAT scores.[16] Within a particular school, the seventy-fifth percentile of the black distribution roughly matches the twenty-fifth percentile of the non-black distribution. These large differences in SAT scores support the essential assumptions made in the models above.[17]

Table 2 examines the distribution of black SAT scores relative to the school average for three tiers of schools. Tier 1 includes those with average SAT scores above 1300, tier 2 between 1200 and 1300, and tier 3 less than 1200. The gap between black SAT scores and the average SAT score

[14] Other cohorts available in the C&B data set include the classes entering in 1951 and 1976. We focus on the 1989 cohort because it is the most recent and more detailed questions were asked about inter-racial interaction.

[15] We experimented with other measures such as private versus public school and found little evidence that similarity in these measures mattered besides through the SAT score.

[16] Note that the SAT has been accused of being biased against blacks. This paper takes no stand on this issue, simply noting that if this is true, it will reflect our empirical results in predictable ways. We return to this in the results section.

[17] Given these differences, our initial thought was that perhaps blacks would look similar to legacies in terms of their SAT scores due to preferential admissions policies for both groups. These preferential admissions policies would then serve as a device that would encourage interaction between blacks and legacies. However, the distribution of legacy SAT scores at particular schools is actually similar to those of non-blacks in general. Note that legacies may still receive preferential treatment in the admissions process. This is because legacies should on average have higher SAT scores than the population of applicants due to their family background. Hence, conditional on being in a particular SAT score range, we would expect legacies to be disproportionately represented at the top end.

Table 1: Distribution of SAT Scores Relative to the School Average[†]

SAT minus School Average SAT	Blacks	Non-Blacks
10th percentile	-340	-153
25th percentile	-261	-67
50th percentile	-170	20
75th percentile	-72	96
90th percentile	9	162
Mean	-166	11
Observations	1561	20566

[†]Sample is taken from the 17 coeducational institutions in the College & Beyond that had participants in the followup survey.

is fairly similar across tiers, though the bottom end of the black distribution is relatively worse at tier 3 schools.

Table 2: Distribution of Black SAT Scores Relative to the School Average By Tier†

Black SAT minus School Average SAT	Tier 1	Tier 2	Tier 3
10th percentile	-302	-344	-360
25th percentile	-233	-247	-281
50th percentile	-143	-154	-191
75th percentile	-59	-72	-87
90th percentile	7	-2	19
Mean	-147	-162	-181
School Average SAT	1331	1251	1123
Observations	5910	7929	8288

† Sample is taken from the 17 coeducational institutions in the College & Beyond that had participants in the followup survey.

4.2 Measures of inter-racial contact

Our key outcome measures are the self-reported answers to questionnaire items inquiring about the friends that respondents made while enrolled in college. For a number of racial and social groups, the survey begins by asking whether respondents knew at least one member of the group. For those responding yes, a follow-up question asks whether the respondent knew two or more group members "well." While these responses are inherently subjective, it is plausible to think that these measures correlate strongly with a latent variable indicating the degree of inter-racial contact respondents experienced while in college. We focus particularly on whether non-blacks knew two or more blacks well and whether whites knew two or more Hispanics or Asians well. Table 3 shows the proportion of individuals answering these questions affirmatively as well as the corresponding shares of blacks, Hispanics, and Asians in the sample. The last three columns disaggregate the numbers by the same tiers as in Table 2.

Representation does appear to correlate with our measure of inter-racial contact. Survey respondents who attended tier 1 schools have the highest probabilities of knowing members from all three groups well and each of the three groups is disproportionately represented at the tier 1 schools. As we move down the tiers, the fraction of both Asians and hispanics fall. This may be related to the location of the college themselves as Hispanics may be more concentrated in particular regions of the country. For example, the only school on the West Coast is a Tier 1 school: Stanford. For blacks, however, we see the same U-shape pattern as in the national data: the middle tier schools have on average the lowest percent black.

Table 3: Probability of Knowing Two of More Members of a Group Well†

	Group	Population	Tier 1	Tier 2	Tier 3
	Blacks	59.8%	67.5%	56.8%	58.0%
Know 2	Asians	59.6%	81.4%	63.3%	48.1%
or More	Hispanics	30.1%	45.7%	38.1%	18.5%
	Blacks	6.4%	7.5%	5.7%	6.2%
Share	Asians	8.2%	12.9%	8.6%	5.6%
	Hispanics	3.3%	5.7%	3.5%	1.9%

† Sample is taken from the 17 coeducational institutions in the College & Beyond that had participants in the followup survey. Tier 1, Tier 2, and Tier 3 refer to institutions with average SAT scores above 1300, between 1200 and 1300, and below 1200 respectively. For knowing two or more blacks well, the sample includes all non-blacks. For knowing two or more Hispanics or Asians, the sample includes only whites. Results are not sensitive to this sample selection.

5 Reduced Form Methods and Results

5.1 Estimation Strategy

We relate the outcome measure, a respondent's self-report of knowing two or more members of a racial or ethnic group well, to observable characteristics of the individual and the school. For our baseline model, we work within the latent dependent variable model framework. Analogously to the latent dependent variable in an ordered response model, the latent variable monotonically related to the number of members of a group k that are well known by an individual i at school j is given by[18]:

$$Y_{ijk}^{*} = \alpha_0 + X_{1i}\alpha_1 + SHR_{jk}\alpha_2 + (SHR_{jk}\alpha_2)^2\alpha_3 + \epsilon_{ijk} \quad (1)$$

where X_{1i} refers to individual-specific characteristics that may affect inter-racial interaction or an individual's standard for knowing someone well. The SHR terms refer to the fraction of the group in question (for example, blacks) at the school. The squared term allows for decreasing (or potentially increasing) returns from increasing representation. Assuming that the error terms are normally distributed and independent from the other variables, we can estimate the probability of knowing two or more well in a probit model where the outcome measure is $Y_{ijk} = 1$ if $Y_{ijk}^{*} \geq \tau_2$, an unknown threshold. All of the α's are allowed to vary across the groups in question.[19]

For a reduced form test of whether similarity in academic background matters, we add to the model the relative SAT score at each school. The underlying index function is then:

$$Y_{ijk}^{*} = \alpha_0 + X_{1i}\alpha_1 + SHR_{jk}\alpha_2 + (SHR_{jk}\alpha_2)^2\alpha_3 + (SAT_i - \overline{SAT}_j)\alpha_4 + \epsilon_{ijk} \quad (2)$$

[18] In the models that follow we are assuming that the latent error term ϵ_{ijk} is distributed independently of the covariates. This assumption may be most likely to be violated regarding the regressor SHR since for historical reasons share black may vary locationally with ways that are correlated with attitudes toward interracial friendship. We discuss below how to control for this potential source of endogeneity bias.

[19] In addition to analyzing inter-racial relationships, we also examine whether the respondent knew well two or more students from three other groups: students from a different region of the U.S., students who were more conservative, and students who were more liberal. We analyze these responses to see whether any relationship we see between test scores and knowing minority students well is driven by social stratification or by those having higher test scores also having different standards for knowing someone well. For example, we expect no relationship between test scores and region of the U.S. conditional on being at the same school. If we then see no relationship between test scores and knowing students from other regions of the country we would conclude that test scores are uncorrelated with the subjective cutoff point for knowing someone well.

If similarity in SAT scores is unimportant— or the SAT score is not related to the characteristic on which similarity matters— α_4 will be zero. Note that it is possible to include school fixed effects in the above specification, with the within-school variation in SAT scores driving the estimate of α_4. Further, because we are using the within-school variation, it is possible to estimate the specification for each school separately and obtain separate estimates of α_4 for each school to further ensure that our estimates not being driven by the features of particular institutions.[20]

We can also explicitly introduce the possibility of a relationship between an individual's own characteristics and the distribution of those same characteristics among the other race. Hence, we develop a model that nests the relationship between the share of group k and interaction rates in equation (1) but allows for similarity in SAT scores across the distribution to matter. In particular, for all i, divide all members of group k into three subgroups: those with SAT scores 160 above their own ($HIGH$), 160 points below their own (LOW), and within 160 points of their own score (MED).[21] Dividing these numbers by the total member of classmates then gives the joint probability of being in the particular SAT group and in the racial group in question. We then allow increasing the shares of each of these groups to differ in their effect on inter-racial interaction. This leads to the following specification:

$$Y^*_{ijk} = \alpha_0 + X_{1i}\alpha_1 + \mathbf{SHR_{jk}}\boldsymbol{\alpha_2} + (\mathbf{SHR_{jk}}\boldsymbol{\alpha_2})^2\alpha_3 + \epsilon_{ijk} \tag{3}$$

where

$$\mathbf{SHR_{jk}}\boldsymbol{\alpha_2} = \frac{\alpha_{20}N_{jkHIGH} + \alpha_{21}N_{jkMED} + \alpha_{22}N_{jkLOW}}{N_j}$$

N_{jkHIGH} refers to the number of students at school j, of race k, who have SAT scores 160 points above individual i while N_j refers to the total number of students at school j. Note that the coefficients on the linear term also enter into the squared term. In this way if one group has little

[20] In this specification, there may be omitted factors correlated with individual SAT scores that influence the rate of inter-personal contact. As a crude example, students at the high end of the SAT distribution for their institution may be "bookworms" who get to know very few other students. To address this concern, we present one specification below that controls simultaneously for relative and absolute SAT scores. Absolute SAT scores prove to be insignificant determinants of inter-racial contact conditional on relative SAT scores. We will also present specifications demonstrating that this is not a universal tendency. White students with high SAT scores are more likely to know members of a group with high average SAT scores – Asians.

[21] 160 points corresponds to the standard deviation in SAT scores across the population of College & Beyond students.

direct benefit then that group will also have little effect on crowd out. That is, if a group has no first-order effect on interaction then the group should also have no second-order effect.

If SAT scores are irrelevant to inter-racial interaction, then the coefficient on the three share variables will be the same. If similarity in SAT scores matters but there is no statistical discrimination, we would expect the coefficient on the share of the group within the SAT range to be higher than the coefficients on the other groups, $\alpha_{21} > \alpha_{20}, \alpha_{22}$. However, we would also expect symmetry between the coefficients on the share of students above and the share of students below, $\alpha_{20} = \alpha_{21}$. If statistical discrimination is present and the group in question has significantly lower SAT scores than the population, the coefficient on the share of the group significantly above the individual's SAT score should be higher than the corresponding coefficient on the share of the group significantly below the individual's SAT score, $\alpha_{20} > \alpha_{22}$.

5.2 Is there evidence that similarity in academic background affects contact rates?

Estimates of a baseline probit of the probability of knowing two or more blacks well are presented in Table 4. The sample consists of all non-black College & Beyond respondents. Variation in percent black is solely at the level of the institution and standard errors are adjusted for clustering. The specification in column 1 gives the results when the only controls are female, percent black, and percent black squared. Females are less likely to report that they know two or more blacks well and this patterns continues throughout the rest of the regression results. However, as we will see later in the paper, this seems to be because females are less likely to report that they know anyone well— what is considered knowing someone well appears to differ across the sexes. Both percent black and percent black squared have the expected signs. With the negative squared term, an interior interaction-maximizing percent black exists at 10%. Only one of the institutions has a student population that is more than 10% black.[22]

[22] The existence of an interior optimum could reflect several different mechanisms. Beyond this point, almost all non-blacks may know two or more blacks well. Thus further increases in representation could increase interaction along a margin we do not directly observe. It is also possible that minority groups begin to eschew cross-racial interaction once their group share reaches some critical mass. We may be observing a "saturation" phenomenon: once representation reaches a critical level, further representation is not necessary to increase interaction. Note that if this is the case, the negative externalities of removing minority students from less-selective campuses where they

The second column adds the individual's SAT score relative to the school average. Those with higher SAT scores at a particular school are substantially less likely to know two or more blacks well. Given that blacks tend to be at the lower end of the within school SAT distribution, this is suggestive that similarities in the characteristics associated with SAT scores may facilitate inter-racial relationships.

It may also be the case, however, that those with high SAT scores are more likely to have discriminatory preferences or are less likely to report that they know someone well. For example, Fryer (2005) finds that those in junior high and high school who have high grades are listed as friends of more students than those who have low grades. Column 3 adds the individual's SAT score to the regression to test whether the result is driven by high SAT score individuals being less likely to interact with blacks. The coefficient on own SAT score was small and insignificant: we cannot reject the exclusion of the individual's SAT score once we control for relative SAT score.[23]

To further test whether similarities in SAT score are important to inter-racial interaction, we perform the same analysis except using the probabilities of knowing two or more Hispanics or Asians well. Results of these specifications, estimated using only the sample of non-Hispanic white students, are reported in table 5. Consistent with the results for knowing blacks, women are less likely to report that they know two or more Hispanics or Asians well. The standards for knowing someone well appear to be different across the sexes.

More important are the effects of relative SAT score. Although about a third of the magnitude of the corresponding coefficient in the black regression, increasing one's SAT score relative to the school average makes interaction with Hispanics less likely. This is consistent with the fact the gap between Hispanic and white SAT scores at the school level is -94 points. The second column displays the results for Asians. Relative SAT score has the opposite effect here: higher relative scores make interaction with Asians more likely. This is consistent with similarity in academic background mattering as Asians on average scored 26 points higher than their white counterparts. Similarity in SAT scores seems to be important not only in inter-racial relationships with blacks,

are well-represented may be minimal. Finally, this point could be the threshold at which the effect of mismatch in backgrounds trumps the effect of increased representation.

[23]We also estimated models where we replaced $(SAT - \overline{SAT})$ with the absolute difference between one's own SAT score and the average SAT score for blacks at their school. We again found strong negative effects of having an SAT score far away from the average SAT score of blacks on the probability of knowing two or more blacks well.

Table 4: Probit Estimates of the Relationship Between Percent Black and Knowing Two or More Blacks Well†

	Specification		
Variable	(1)	(2)	(3)
Share Black×10	1.500	1.507	1.248
	(0.293)	(0.294)	(0.609)
(Share Black×10)^2	-0.766	-0.752	-0.581
	(0.204)	(0.294)	(0.609)
Female	-0.088	-0.140	-0.140
	(0.039)	(0.040)	(0.040)
$(SAT - \overline{SAT})$ (00's)		-0.105	-0.124
		(0.017)	(0.047)
SAT			0.019
			(0.024)
Constant	-0.262	-0.230	-0.438
	(0.099)	(0.010)	(0.407)
Log likelihood	-4756	-4721	-4721

†N=7126. Sample is all non-blacks who had valid answers for the questions on knowing two or more blacks well in the 1996 followup survey.

but also in inter-racial relationships with Hispanics and Asians.

Table 5: Probit Estimates of the Relationship Between Group Share and the Probability Knowing Two or More Members of the Group Well†

	Know Two or More:	
Variable	Hispanics	Asians
Group Share×10	2.954	1.646
	(0.393)	(0.259)
(Group Share×10)^2	-1.043	-0.331
	(0.451)	(0.148)
Female	-0.141	-0.237
	(0.025)	(0.046)
$(SAT - \overline{SAT})$ (00's)	-0.038	0.075
	(0.013)	(0.021)
Constant	-1.244	-0.621
	(0.056)	(0.091)
Log likelihood	-3073	-3296
Average Group SAT - Average White SAT	-94	26

† N=5557. Sample is all non-Hispanic whites who had valid answers for the questions on knowing two or more blacks well in the 1996 followup survey.

Although including school fixed effects makes it impossible to identify how representation affects interaction, we can see if the coefficient on relative SAT score changes once school fixed effects are included. Results of these specifications are in Table 6. Once again we see that those with high relative SAT scores are more likely to interact with Asians and less likely to interact with blacks and Hispanics. In addition to the results in Table 6, we estimated the model with major effects and with school-major effects. In all cases the same results emerged with little change in the magnitudes

Table 6: Probit Estimates of the Relationship Between Group Share and the Probability Knowing Two or More Members of the Group Well with School Fixed Effects[†]

	Know Two or More:					
				Same	More	More
Variable	Blacks	Hispanics	Asians	Region	Conservative	Liberal
Female	-0.137	-0.157	-0.249	-0.190	-0.115	-0.316
	(0.060)	(0.023)	(0.044)	(0.068)	(0.036)	(0.045)
$(SAT - \overline{SAT})$ (00's)	-0.105	-0.043	0.075	0.008	0.025	-0.014
	(0.015)	(0.012)	(0.022)	(0.031)	(0.015)	(0.024)
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

[†] For blacks, sample is all non-whites who had valid answers for the questions on knowing two or more blacks well in the 1996 followup survey N=7126. For Hispanics and Asians the corresponding sample is for whites only N=5547. For same region, more conservative and more liberal, we used responses by non-blacks N=7428, 7262, and 7232 with sample variation due to different skip patterns. Results are not sensitive to the sample specification.

and no change it the statistical significance of the relative SAT score.

The last three columns repeat the analysis but with the dependent variable now being knowing two or more students well from different regions of the U.S., who are more conservative than the respondent, and who are more liberal than the respondent. We had no prior reason to believe that there was any relationship between test scores and region of the country or political leanings. In all three cases, the effect of test scores[24] on contact was small and insignificant. This suggests that the relationship we see between test scores and inter-racial contact is driven by sorting rather than by some relationship between test scores and standards for knowing someone well. In contrast, the negative and significant coefficient estimates on female in these three cases indicate that the negative relationship between female and knowing two or more minorities well is due to females having different standards for knowing someone well.

Finally, since we are relying on within school variation in SAT scores, we can estimate the model separately for each institution. Again we see strong evidence that similarity in academic backgrounds matter. For twelve of the seventeen institutions the effect of an individual's SAT score on the probability of knowing two or more blacks well was negative and significant. In only one case was there a positive sign but even then the coefficient was not close to significant. In seven of the seventeen institutions an individual's SAT score had a positive and significant of knowing two or more Asians well. While five of the institutions had negative signs, none of the five were significant. Results for Hispanics were generally insignificant.

5.3 Further Tests of Likes Interacting with Likes and Evidence of Statistical Discrimination

With evidence to this point suggesting that those with high SAT scores at a particular school are less likely to know two or more blacks well, we now attempt to relate this back to the distributions of SAT scores of blacks and non-blacks within the school. In particular, we create three individual-specific peer groups: those with SAT scores 160 points above the individual's own SAT score, 160 points below, and those with SAT scores within plus or minus 160 points. We then examine how percent black within these groups affects the probability of knowing two or more blacks well. The

[24] By conditioning on institution effects, it is not possible to distinguish between the effects of relative and absolute test scores.

shares are adjusted for the size of the group such that the initial model is nested: if the coefficients of percent black in the three groups are the same, we will obtain the results from the initial model. Results are displayed in Table 7.

The first column of Table 7 uses the institutional variation used in the previous tables. Similarity in SAT scores seems extremely important to inter-racial interaction. The coefficient on the black share of those of similar SAT scores is four times as large as the black share of those with SAT scores significantly below the individual's score.

The response to share black in the different peer groups, however, is not symmetric. Namely, the coefficient on black share for those with significantly higher SAT scores is actually higher than corresponding coefficient on similar SAT scores. To test whether this difference is statistically significant, the second column of the table restricts the effect of the higher SAT group share to be the same at those with similar SAT scores. We cannot reject the restriction that the coefficients are the same.

That the coefficients are the same for higher SAT score share and similar SAT score share stands in contrast to the significantly smaller effects for the lower SAT score share. This is consistent with a model of statistical discrimination where individuals only receive signals about their compatibility. Individuals would prefer to match with those of similar academic backgrounds, but must choose to engage in relationships based only on signals of academic backgrounds. High SAT score blacks emit signals that are weighted down by the population average SAT score for blacks. A white individual may be more likely to explore a relationship with a black individual with a significantly higher SAT score than one with a similar SAT score to his own due to statistical discrimination. The probability of the relationship being successful, however, is higher with the one who has a similar SAT score. The estimates suggest that these two effects–exploring the relationship and having the relationship succeed–balance out when considering relationships with blacks whose SAT scores are the same or similar to one's own.

The next set of columns expands the number of groups to six. Each of the SAT groups is split in two, one for within the individual's major and one outside. Again if the coefficient estimates are the same we are back to the restricted model. Although the estimates are much more noisy, the same patterns emerge both within and outside one's major. Regardless of whether it is outside or inside the major, the black share for those with lower SAT scores is much smaller than the corresponding effects through the black shares of higher SAT scores and similar SAT scores.

Note that the estimates in this section do not control for individual SAT score. Rather, estimation in this section is designed to put a more structural interpretation behind the negative relationship between SAT score and interaction with blacks. Adding SAT score to the regressions yields noisier estimates and we can no longer distinguish between the coefficients on the different share variables. This is not surprising given that we only have rough groupings for how similar individuals are and SAT score will be correlated with the true underlying groupings. The argument for keeping SAT score out of these regressions is that we do not believe the SAT score itself is causal to whether one knows two or more blacks well, particularly given the relationship between SAT score and knowing members of other groups well, such as Asians, Hispanics, those with different political leanings, and region of the country. We are instead providing a mechanism for the reduced form result.

6 Structural Model and Results

In this section we specify a model of interactions from which we can perform the policy simulations. We first describe how potential friends arrive and how individuals decide whether or not to form a friendship. We next describe our simulated maximum likelihood procedure where simulation is necessary because we do not actually observe whether a friendship is formed. The limited data on interactions means that some of our parameters are not identified. However, conditional on setting one of the parameters in the model we are able to identify the others. We estimate the model and show how the results change given different values for the unidentified parameter.

6.1 Model

With the reduced form evidence suggesting that similarity in academic background affects interracial contact and that asymmetries may be present, we now to proceed to the formulation and estimation of the structural model. With the structural model we will be able to forecast how interactions will change both from the changing composition of the student body and through the corresponding updating of beliefs about the academic backgrounds of their peers. All students are assigned N potential friends from among their student bodies. The probability of an individual

Table 7: Probit Estimates of the Relationship Between Percent Black Both In and Out of Major and Knowing Two or More Blacks Well Allowing for Variation by SAT Gap†

		Specification					
	Variable	(7)	(8)	(9)	(10)	(11)	(12)
Black Share	Higher SAT	17.66	—				
Overall		(6.47)					
	Similar SAT	14.98	15.41				
		(2.55)	(2.35)				
	Lower SAT	3.65	3.63				
		(1.24)	(1.26)				
Black Share	Higher SAT			7.40	—	34.91	—
in Major				(49.17)		(51.95)	
	Similar SAT			26.66	23.8	37.82	36.57
				(15.08)	(11.80)	(18.13)	(13.60)
	Lower SAT			1.91	1.91	-1.95	-1.81
				(6.69)	(6.57)	(7.38)	(7.31)
Black Share	Higher SAT			19.56	—	15.46	—
Outside Major				(9.60)		(9.23)	
	Similar SAT			13.96	14.62	12.09	12.66
				(2.71)	(2.49)	(2.78)	(2.50)
	Lower SAT			4.01	3.93	4.43	4.37
				(1.57)	(1.54)	(1.70)	(1.67)
	Share Black^2	-0.314	-0.315	-0.316	-0.315	-0.335	-0.336
		(0.032)	(0.032)	(0.031)	(0.031)	(0.035)	(0.035)
	Female	-0.141	-0.140	-0.138	-0.138	-0.161	-0.161
		(0.031)	(0.031)	(0.031)	(0.031)	(0.032)	(0.032)
	Constant	-0.090	-0.093	-0.100	-0.100	-0.220	-0.222
		(0.064)	(0.064)	(0.065)	(0.064)	(0.101)	(0.100)
	Major Fixed Effects	No	No	No	No	Yes	Yes
	Log likelihood	4729.9	4730.0	4729.6	4729.7	4716.9	4717.0

†N=7126. Sample is all non-blacks who had valid answers for the questions on knowing two or more blacks well in the 1996 followup survey.

being assigned n potential friends of group j follows a binomial distribution:

$$Pr(n) = \begin{pmatrix} N \\ n \end{pmatrix} \lambda_j^n (1-\lambda_j)^{N-n} \quad n = 0, 1, ...N$$

where λ_j is given by:

$$\lambda_j = \lambda_0 + \lambda_1 SHR_j$$

One student among the pair is exogenously assigned to make the decision as to whether the friendship forms given the information the student has at the time. Match quality from the perspective of the decision maker is a function of the potential friend's academic background, A, which the decision-maker only observes with noise, X, the observed characteristics of the decision-maker, and ϵ, an unobserved match-specific component. Decision-makers know the distribution of A conditional on j and observe a signal on A as well. This signal, S, is given by:

$$S = A + \zeta$$

where ζ is distributed N$(0, \sigma)$. A is assumed to be discrete, taking on one of K values. The probability of A taking on the value A_k conditional on the partner being a member of j is π_{jk}. The decision-maker updates his beliefs about A using Bayes' Rule:

$$P(A_k|S,j) = \frac{\pi_{jk}\mathcal{L}(A_k|S)}{\sum_{k'=1}^{K} \pi_{jk'}\mathcal{L}(A_{k'}|S)} \tag{4}$$

The gains from matching with someone of match quality A_k depend upon how similar A_k is to the individual's own academic background. In practice, we measure this similarity by the absolute distance between the two academic backgrounds. The expected utility for student i from accepting the nth match with a member of group j is then given by:

$$EU_{in} = \frac{\sum_{k=1}^{K} \pi_{jk}\mathcal{L}(A_k|S_n)(\alpha_0|A_i - A_k|)}{\sum_{k'=1}^{K} \pi_{jk'}\mathcal{L}(A_{k'}|S_n)} + X_n\alpha_1 + \epsilon_{in} \tag{5}$$

where ϵ_{in} is known to the individual but not the econometrician. Student i accepts the match when $EU_{in} > 0$. We assume that the ϵ's are iid Type I extreme value, implying that the probability of a match being accepted from the perspective of the econometrician is:

$$p_{in} = \frac{\exp{(EU_{in} - \epsilon_{in})}}{\exp{(EU_{in} - \epsilon_{in})} + 1} \tag{6}$$

In practice we use SAT scores as our approximation of A. The empirical distribution of SAT scores for group j is then used in formulating the π_{jk}'s. The parameters to be estimated are then the arrival rate parameters, the λ's, the utility function parameters, the α's, and the variance on the signal of academic background, σ. The next section describes the procedure used to estimate these parameters.

6.2 Simulated Maximum Likelihood Procedure

In the College & Beyond dataset, we observe very little about the matching process. We do not know anything about same-race friendships nor do we know the number of friends–actual or potential– a student has. The only information we have is whether a student had two more friends of particular ethnic group that was not their own ethnic group. We know nothing about the characteristics of friends from other races. This necessarily restricts the scope of estimation. We focus our attention on the decisions by non-blacks as to whether to accept black matches.

We use simulated maximum likelihood as a means of obtaining consistent parameter estimates given the limited available data. A simulated draw involves three steps:

1. Drawing the set of potential black matches for each individual,
2. Drawing the signals for each potential match,
3. Given the signals and the set of potential matches, forming the probability that the individual has two or more black friends.

We do this R times for each individual and then average over the R simulations. Each of these steps is described in more detail below.

6.2.1 Drawing the set of potential matches

Given the arrival rate parameters λ_0 and λ_1, we could calculate the probability of having any number of potential black matches directly from the binomial formula. However, the likelihood will not be smooth if every time the arrival parameters are updated we took new draws on the number of potential matches.

Instead, we simulate the number of potential black matches by drawing an integer between 0 and 8 from the discrete uniform distribution. If an individual is assigned eight potential black friends, it is almost certain they will accept two or more of the friendships. Hence, eight actually refers to being assigned eight or more potential black friends. We do this for each individual R times.

Label the number of potential black friends for rth draw for individual i as $N_{bi}^{(r)}$ Then, when the likelihood is maximized, we calculate the actual probabilities of being assigned $N_{bi}^{(r)}$ friends and use these as weights when we average across the R simulations. For example, suppose $N_{bi}^{(r)} = 0$. The sampling procedure would yield this outcome 1/9 of the time. Within the maximization routine, if the binomial formula placed a probability lower than 1/9 of drawing zero black friends, then the rth draw would be down-weighted for this individual. The exact weights are defined in section 6.2.3.

6.2.2 Drawing the signals for each potential match

Conditional on being assigned $N_{bi}^{(r)}$ friends, we then randomly draw $N_{bi}^{(r)}$ individuals from the black population at the college i attended. We then need to draw the signals that dictate whether a match is accepted. We do this by drawing the signals from a standard normal distribution, again outside of the maximization routine. Then, these signals are weighted up or down depending upon the current value of σ. With this draw, we can compute the perceived probability that the individual is of the kth ability type, $P(A_k|S, j)$, using (4).

6.2.3 Forming the likelihood

Conditional on the signal probabilities and the individual's own characteristics, X_i, we can form the probability that each of the $N_{bi}^{(r)}$ matches is accepted using the expression in (6). Label the probability of the $n_i^{(r)} \in \left\{0, ..., N_{bi}^{(r)}\right\}$ match being accepted as $p_{in}^{(r)}$. Next note that from these match probabilities for draw r, we can compute the probability of two or more matches, $p_{2i}^{(r)}$, using the binomial formula:

$$p_{2i}^{(r)} = 1 - \prod_{n=1}^{N_{bi}^{(r)}} \left(1 - \lambda p_{in}^{(r)}\right) - \sum_{n=1}^{N_{bi}^{(r)}} \lambda p_{in}^{(r)} \prod_{n' \neq n}^{N_{bi}^{(r)}} \left(1 - \lambda p_{in'}^{(r)}\right) \tag{7}$$

Equation (7) gives the conditional probability of two or more matches, where we are conditioning on a particular simulation draw. We now need to weight the simulation draws according to the estimates of λ. In particular, denote $w_i^{(r)}$ as the binomial probability associated with draw r:

$$w_i^{(r)} = \left(\begin{array}{c} N \\ N_{bi}^{(r)} \end{array} \right) \lambda_j^{N_{bi}^{(r)}} (1 - \lambda_j)^{N - N_{bi}^{(r)}}$$

The total weight assigned to draw r for individual i, $W_i^{(r)}$, is then given by weight associated with r divided by the sum of the R weights:

$$W_i^{(r)} = \frac{w_i^{(r)}}{\sum_{r'=1}^{R} w_i^{(r')}}$$

Our simulation estimator of the unconditional probability of knowing two or more blacks well is then:

$$\hat{p}_{2i} = \sum_{r=1}^{R} W_i^{(r)} p_{2i}^{(r)}$$

Noting that this simulated probability is a function of the unknown parameters, collected and denoted here by θ, we can now construct a maximum likelihood estimator of $\theta = (\sigma, \gamma, \alpha)$. Let d_i denote whether individual i had 2 or more black friends. We then pick the value of the unknown parameters to maximize the simulated likelihood function:

$$\hat{\theta} = \arg\max_{\theta \in \Theta} \frac{1}{\mathcal{N}} \sum_{i=1}^{\mathcal{N}} d_i \log \hat{p}_{2i}(\theta) + (1 - d_i) \log(1 - \hat{p}_{2i}(\theta))$$

Letting the number of simulation draws increase at a rate faster than the square root of the sample size, the simulated MLE will be asymptotically equivalent to the regular MLE- see, e.g. Hajivassiliou and Ruud (1994).

The above discussion assumes that N, the number of potential friends as opposed to the potential number of black friends (N_{bi}), is known . Separate identification of N and λ is not possible except through functional form restrictions. We therefore estimate the model for different values of N and note how N affects the results. The policy simulations are also conducted at different values of N.

6.3 Estimates

Estimates of the model for $N = 40$ are displayed in Table 8. the results for other values of N between 20 and 100 are given in the appendix. The first two rows give the arrival parameters. As can be seen from the appendix, these are the parameters that are sensitive to the choice of N. Increasing N leads to smaller values for both λ_0 and λ_1 as the probability of meeting black individuals is increasing in N, λ_0, and λ_1.

While the arrival rate parameters vary with N, the utility function parameters remain virtually unchanged as N changes. The reason for this is that these coefficients are driven by differences in interaction rates across individuals of different academic backgrounds–changing N does not affect the relationship between academic background and interaction. Most important to understanding the tradeoff between representation and similarity in academic background is the estimate of the parameter on the difference between own SAT score and the SAT score of the partner. Here we see that similarity in academic background is significant in determining successful matches. Since the model is highly non-linear–particularly given individuals only have expectation on academic backgrounds– how big of an effect this is is difficult to see from the estimates themselves. In the next section we use different rules for assigning students to schools which will make the tradeoff explicit. The estimate of the variance on the signal of academic background suggests– consistent with the reduced form evidence– some scope for statistical discrimination.[25]

Table 8: Estimates of the Structural Model†

	Coefficient	Standard Error
Arrival Parameters		
λ_0	0.0278	0.0034
λ_1	0.2071	0.0278
Utility Parameters		
Intercept	2.7314	0.6673
Sex_i	-0.4015	0.1383
$\lvert A_n - A_i \rvert$ (000's)	-3.9506	1.0296
σ	0.3595	0.2071
Log likelihood	-4746	

†N=7126. Dependent variable is the probability of knowing two or more blacks well.

[25] Since this coefficient must be greater than or equal to zero, the test for statistical significance is a one-tail test. The t-stat is 1.74, making the coefficient significant at the 95% level.

7 Alternative Assignment Rules

The evidence presented in the previous sections suggests that rates of inter-racial contact are governed both by minority representation and the degree of background similarity between students of different races. Given the large differences in SAT scores across blacks and whites at the College & Beyond schools and the locally positive relationship between college quality and percent black within highly selective colleges, racial preferences may actually have a negative effect on inter-racial interaction. While we do not have the data to examine this claim directly, we are in a position to analyze the potential impact of altering the assignment of blacks and non-blacks across the institutions represented by the College & Beyond.

To do this, we need to quantify the relationship between race and college assignment. We estimate a multinomial logit model of college assignment where the probability of being assigned to one of the j schools in the data set is a function of the individual's race and SAT score. The probability of being assigned to school j is given by:

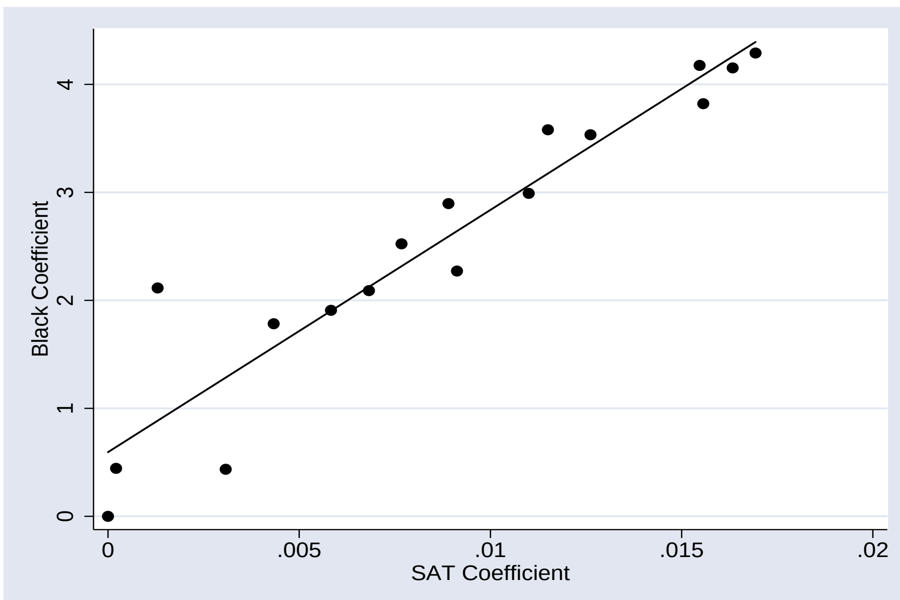
$$Pr(j|X_i) = \frac{\exp(\beta_{0j} + BLACK_i\beta_{1j} + SAT_i\beta_{2j})}{\sum_{j'=1}^{J}\exp(\beta_{0j'} + BLACK_i\beta_{1j'} + SAT_i\beta_{2j'})}$$

Figure 3 plots the relationship between the coefficient on SAT and the coefficient on black. The coefficient on black is highest when the coefficient on SAT score is highest. The R^2 of the fitted line is 0.89. Consider two individuals, one of whom is black. The points on the fitted line imply that the non-black individual would need to have an SAT score 224 points higher than his black counterpart in order to face the same assignment rules.

We examine assignment policies that attempt to break the tie between SAT scores of the school and the race of the individual. Given the coefficients associated with the fitted line, we can simulate assignment rules that are less aggressive than those used in the College & Beyond. Namely, we can rotate this line around the mean of the SAT coefficients. A less aggressive assignment policy would then be associated with a flattening of this line.

The residuals from the fitted line represent the advantages particular schools have in attracting black students. However, there are not enough schools to be representative of the population of schools in each tier. Hence, all simulations are done using points on this line— purging the school-specific residuals from the black coefficient. Note further that some individuals who are in the College & Beyond data set would not have been able to attend any of those schools without race preferences – the data simply does not contain any completely unselective colleges. Our procedure will, admittedly unrealistically, assign the least qualified students of either race to the least selective college in the sample, even though in reality those students might attend non-sampled unselective colleges in the presence of a policy change. While we have no way of evaluating the full effects of removing race conscious admissions, we can partially address this issue by omitting the least selective sampled institution in the policy analysis.[26] To the extent that our procedure leads minority students to "pool" at the least selective institution, this correction will improve our estimate of the overall impact. The least selective sampled institution thus becomes the representative of all unselective colleges.

Figure 3: The Relationship Between the Coefficients on Black and SAT in the School Assignment Policies†



†Coefficients from a multinomial logit model of college assignment normalized with respect to the institution with the lowest coefficient on black. See text for details.

The estimated effects of dampening the relationship between the SAT coefficients and the black coefficients on the probability of knowing two or more blacks well are displayed in Table 9. The

[26]This ignores the increased interaction at lower-tiered schools that would result from weakening the current rules. It also ignores the fact that some colleges below those in the data also practice affirmative action. We cannot address either of these with our data.

first column shows the aggregate effect while the next set of columns breaks out the effects by tier. Assigning blacks according to the white rules would actually yield a small increase in the probability of knowing two of more whites well. The slope of the black-SAT coefficient line that maximizes inter-racial contact is 80% of the actual slope. In other words, a small degree of racial admissions preferences appears to maximize the population rate of inter-racial contact, but the observed degree of preferences goes well beyond this maximum.

It is interesting to also consider how share black would vary across the tiers under the different assignment rules. Under the current assignment rule, tier 1 schools are the most diverse. Weakening the relationship between black and the assignment rules by twenty percent leads to the most even distribution across the tiers. Assigning blacks according to the white rules does lead to large differences in the share black across schools. Namely, share black would be more than four times higher at tier 3 schools than at tier 1 and more than twice as high at tier 3 schools than at tier 2 schools.

Clearly as racial preferences are weakened some blacks would be unable to obtain admission to any of the institutions in the College & Beyond. To take this into account, we repeated the assignment rules using the full set of schools and then removed the least selective sampled school when discussing the policy effects. The simulations would then be representative of what would actually happen with the removal of racial preferences if the shift of blacks to the least selective school is representative of the shift that would occur out of the College & Beyond data set. Removing the least selective school from the policy analysis did show that completely removing racial preferences from the rest of the College & Beyond schools would lead to less inter-racial contact. However, the interaction maximizing policy would still involve a 50% reduction of the slope of the black-SAT coefficient line.

The small changes at the aggregate level, however, are masking much larger changes across the tiers of institutions. Namely, top tier schools see large drops in interaction rates as race preferences are weakened with corresponding larger increases in interaction rates at tier 3 schools. A social welfare function which puts a higher premium on interactions at top tier schools may then find that the base assignment rules are optimal.

The effects of varying the assignment rules on the probability of knowing two or more blacks well are also not uniform within a school. In Table 10 we examine how this probability changes

Table 9: Probabilities of Knowing Two or More Blacks Well Under Various Assignment Rules†

	Change in Interaction			
	Overall	Tier 1	Tier 2	Tier 3
10% weaker	0.34%	-1.39%	0.13%	1.39%
20% weaker	0.62%	-2.79%	0.15%	2.73%
30% weaker	0.85%	-4.18%	0.06%	4.04%
40% weaker	1.03%	-5.56%	-0.10%	5.29%
50% weaker	1.18%	-6.88%	-0.37%	6.51%
60% weaker	1.28%	-8.16%	-0.73%	7.64%
70% weaker	1.34%	-9.39%	-1.16%	8.71%
80% weaker	1.36%	-10.54%	-1.67%	9.72%
90% weaker	1.35%	-11.59%	-2.23%	10.65%
100% weaker	1.32%	-12.57%	-2.85%	11.50%

†See text for details on how the simulations were conducted.

for those who have the highest test scores within their school and those who have the lowest. The standard deviation of SAT minus school average SAT for non-blacks is 125 in our data. We then examine the effects of changing the assignment rules for those non-blacks with scores 125 points above or below the school average.

Overall, those with the highest scores within a school see their interaction rates increase the most as the assignment rules for blacks become more like the assignment rules for whites. Indeed, for this group interaction is maximized at the corner where blacks are assigned according the white rules. Again, the benefits are coming primarily from tier 3 schools at the expense of the other schools. This is true for those with the lowest scores at each school as well. However, the blacks that are pushed from the top schools lower down have a bigger effect on those at the top SAT distribution within a school than on those at the bottom.

The results for Tier 2 schools illustrate the heterogeneity in the effects of weakening the tie between the black and SAT score coefficients. The lowest interaction rates for top SAT score students at tier 2 schools are observed under the current assignment policy. However, the current assignment policy leads to the highest interaction rates for those at tier 2 schools with the lowest SAT scores.

8 Conclusion

It is commonly argued that increasing minority representation on selective college campuses will increase the likelihood, frequency, and intensity of inter-racial interaction at those institutions. While we offer no evidence here to directly contradict this presumption, this paper makes two counterpoints rooted in basic economic models of friendly interaction between agents. The first point is exceedingly straightforward. Policies that influence only the distribution of minority students, and not their total number, necessarily reduce representation on some campuses. In practice, this reduction tends to occur at moderately-selective institutions. Even in a simple model of random interaction between agents on a campus, the net impact of racial preferences in admissions can only be considered positive if interaction at these institutions is valued less than interaction at the most selective institutions. One might worry little about this effect if minorities were initially well-represented at moderately-selective institution, because the marginal impact of removing a minority student on the margin may be low when a large number remain. In practice, however,

Table 10: Probabilities of Knowing Two or More Blacks Well Under Various Assignment Rules Conditional on SAT Rank†

	Overall		Tier 1		Tier 2		Tier 3	
	+ 1 std.	-1 std.	+ 1 std.	-1 std.	+ 1 std.	-1 std.	+ 1 std.	-1 std.
10% weaker	0.92%	0.04%	-1.29%	-1.73%	0.58%	-0.30%	1.92%	0.93%
20% weaker	1.76%	0.04%	-2.59%	-3.47%	0.94%	-0.70%	3.79%	1.83%
30% weaker	2.55%	-0.02%	-3.87%	-5.20%	1.23%	-1.20%	5.63%	2.68%
40% weaker	3.27%	-0.09 %	-5.15%	-6.90%	1.47%	-1.75%	7.34%	3.52%
50% weaker	3.95%	-0.02%	-6.38%	-8.55%	1.52%	-2.42%	9.05%	4.34%
60% weaker	4.55%	-0.35%	-7.54%	-10.14%	1.46%	-3.14%	10.66%	5.08%
70% weaker	5.11%	-0.52%	-8.63%	-11.68%	1.37%	-3.93%	12.17%	5.80%
80% weaker	5.61%	-0.73%	-9.67%	-13.13%	1.11%	-4.80%	13.65%	6.44%
90% weaker	6.05%	-0.96%	-10.55%	-14.45%	0.83%	-5.69%	14.95%	7.01%
100% weaker	6.43%	-1.18%	-11.40%	-15.72%	0.47%	-6.62%	16.18%	7.59%

†See text for details on how the simulations were conducted.

moderately-selective colleges have fewer minority students than highly-selective colleges.

The second point follows from a long history of social scientific research establishing that individuals tend to associate with those who are similar to them along a number of dimensions. This observation motivates a more compelling model of interpersonal interaction, where agents actively choose whether to invest in friendships. Empirically, we show that the probability of interaction between races on a campus is sensitive to the degree of mismatch between racial groups, as measured by SAT scores.[27] This sensitivity may lead to statistical discrimination or to behavioral norms within a school that are more race-specific than they otherwise would be.[28]

Because race preferences serves to exacerbate these mismatches, the increase in inter-racial contact associated with its practice is at best weakly positive. Our simulations indicate that racial admissions policies as currently practiced actually have a mild negative impact on population rates of inter-racial contact. More generally, results indicate that the use of racial preferences has only a very small impact on the population rate of inter-racial contact. Once again, in light of this evidence the racial preferences can be considered positive only if interaction at lower-tiered schools is valued less than that at the most selective universities.

We intend for these results to illustrate trade-offs present when admissions officers utilize racial preferences in admissions. Our results do suggest that racial preferences are a relatively weak tool for increasing the amount of between-race interaction that takes place on a particular campus, and, because diversity at one institution precludes diversity at another, an even weaker tool for increasing the total amount of such interaction in the population.

[27]Note that mismatch is different here from the standard argument that racial preferences put minorities in schools where they can not succeed. See Rothstein and Yoon (2006) for empirical tests of this argument.

[28]See Akerlof and Kranton (2000) for a discussion of identity in economics.

Appendix

A Proofs

A.1 Proof of Proposition 1

Proof. The selective university solves:

$$\max_{\pi_{WH1},\pi_{BL1}} \quad \theta_1(\pi_{BH}+\pi_{WH1})+\pi_{WH1}(\pi_{BH}+\gamma\pi_{BL1})$$
$$\text{s.t. } \pi_{BH}+\pi_{BL1}+\pi_{WH1}=1$$

Substituting in the constraint for π_{WH1} yields

$$\max_{\pi_{BL1}} \quad \theta_1(1-\pi_{BL1})+(1-\pi_{BH}-\pi_{BL1})(\pi_{BH}+\gamma\pi_{BL1})$$

Taking the derivative with respect to π_{BL1}

$$\frac{d}{d\pi_{BL1}}=0=-\theta_1-\pi_{BH}+\gamma(1-\pi_{BH}-2\pi_{BL1})$$

Solving for π_{BL1}

$$\pi_{BL1}=-\frac{\theta_1}{2\gamma}-\frac{(1+\gamma)\pi_{BH}}{2\gamma}+\frac{1}{2}$$

We are looking for conditions under which

$$\pi_{BL1}+\pi_{BH}>\frac{\pi_B}{2}$$

Note that

$$\frac{(1+\gamma)\pi_{BH}}{2\gamma}=\left(\frac{1}{2\gamma}+\frac{1}{2}\right)\pi_{BH}=\pi_{BH}+\frac{(1-\gamma)\pi_{BH}}{2\gamma}$$

Therefore,

$$\pi_{BL1}+\pi_{BH}=\pi_{B1}=-\frac{\theta_1}{2\gamma}-\frac{(1-\gamma)\pi_{BH}}{2\gamma}+\frac{1}{2}$$

We want to know when this will be $>\frac{\pi_B}{2}$

$$\begin{aligned}
\tfrac{1}{2}-\tfrac{\theta_1}{2\gamma}-\tfrac{(1-\gamma)\pi_{BH}}{2\gamma} &> \frac{\pi_B}{2}\\
1-\tfrac{\theta_1}{\gamma}-\tfrac{(1-\gamma)\pi_{BH}}{\gamma} &> \pi_B\\
\gamma-\theta_1-(1-\gamma)\pi_{BH} &> \gamma\pi_{BL}+\gamma\pi_{BH}\\
\gamma-\theta_1-\pi_{BH}-\gamma\pi_{BL} &> 0
\end{aligned}$$

□

A.2 Proof of Proposition 2

Proof. Given that B types are a minority (assumption 1) and that the school could fill up using only H types (assumption 5) it is clear that the selective university will choose to admit all of the B types of academic background H and none of the W types of academic background L: $\pi_{BH1} = \pi_{BH}$, $\pi_{WL2} = \pi_{WL}$. Taking this as given, maximizing inter-racial contact solves:

$$\begin{aligned} \max_{\pi_{WH1},\pi_{BL1}} \quad & \pi_{WH1}(\pi_{BH} + \gamma\pi_{BL1}) + [\gamma(\pi_{WH} - \pi_{WH1}) + \pi_{WL}](\pi_{BL} - \pi_{BL1}) \\ & \text{s.t. } \pi_{WH1} + \pi_{BL1} + \pi_{BH} = 1 \end{aligned}$$

Substituting in the constraint for π_{WH1} yields

$$\max_{\pi_{BL1}} \quad (1 - \pi_{BH} - \pi_{BL1})(\pi_{BH} + \gamma\pi_{BL1}) + [\gamma(\pi_{WH} - 1 + \pi_{BL1} + \pi_{BH}) + \pi_{WL}](\pi_{BL} - \pi_{BL1})$$

Taking the derivative with respect to π_{BL1}

$$\frac{d}{d\pi_{BL1}} = 0 = -\pi_{BH} + \gamma(1 - \pi_{BH} - 2\pi_{BL1}) + \gamma\pi_{BL} - \gamma(\pi_{WH} - 1 + 2\pi_{BL1} + \pi_{BH}) - \pi_{WL}$$

Combining terms yields

$$0 = 2\gamma - (1 + 2\gamma)\pi_{BH} - 4\gamma\pi_{BL1} + \gamma\pi_{BL} - \gamma\pi_{WH} - \pi_{WL}$$

Solving for $\pi_{BL1} + \pi_{BH}$

$$\begin{aligned} \pi_{BL1} &= \frac{1}{2} - \frac{(1 + 2\gamma)\pi_{BH}}{4\gamma} + \frac{\pi_{BL}}{4} - \frac{\pi_{WH}}{4} - \frac{\pi_{WL}}{4\gamma} \\ \pi_{BL1} + \pi_{BH} &= \frac{1}{2} - \frac{(1 - 2\gamma)\pi_{BH}}{4\gamma} + \frac{\pi_{BL}}{4} - \frac{\pi_{WH}}{4} - \frac{\pi_{WL}}{4\gamma} \qquad (8) \end{aligned}$$

We want to show that

$$\pi_{BL1} + \pi_{BH} < \frac{\pi_{BL} + \pi_{BH}}{2}$$

Therefore,

$$\begin{aligned} \frac{1}{2} - \frac{(1 - 2\gamma)\pi_{BH}}{4\gamma} + \frac{\pi_{BL}}{4} - \frac{\pi_{WH}}{4} - \frac{\pi_{WL}}{4\gamma} &< \frac{\pi_{BL} + \pi_{BH}}{2} \\ \frac{1}{2} - \frac{(1 - \gamma)\pi_{BH}}{4\gamma} - \frac{\pi_{WH}}{4} - \frac{\pi_{WL}}{4\gamma} &< \frac{\pi_{BL} + \pi_{BH}}{4} \end{aligned}$$

$$
\begin{aligned}
2-\frac{(1-\gamma)\pi_{BH}}{\gamma}-\pi_{WH}-\frac{\pi_{WL}}{\gamma} &< \pi_{BL}+\pi_{BH} \\
2\gamma-(1-\gamma)\pi_{BH}-\gamma\pi_{WH}-\pi_{WL}-\gamma\pi_{BL}-\gamma\pi_{BH} &< 0 \\
2\gamma-(1-\gamma)\pi_{BH}-(1-\gamma)\pi_{WL}-\gamma\underbrace{(\pi_{WH}+\pi_{WL}+\pi_{BL}+\pi_{BH})}_{=2} &< 0 \\
\underbrace{-(1-\gamma)\pi_{BH}-(1-\gamma)\pi_{WL}}_{(-)} &< 0 \qquad \square
\end{aligned}
$$

A.3 Proof of Proposition 3

Proof. The social planner maximizing the selective university's objective function plus the interaction at school 2 can be rewritten to solve:

$$
\max_{\pi_{WH1},\pi_{BL1}} \quad (\theta_1-\theta_2)(\pi_{BH}+\pi_{WH1})+\pi_{WH1}(\pi_{BH}+\gamma\pi_{BL1})+[\gamma(\pi_{WH}-\pi_{WH1})+\pi_{WL}](\pi_{BL}-\pi_{BL1})
$$
$$
\text{s.t. } \pi_{WH1}+\pi_{BL1}+\pi_{BH}=1
$$

Substituting in the constraint for π_{WH1} yields

$$
\begin{aligned}
\max_{\pi_{BL1}} \quad & (\theta_1-\theta_2)(1-\pi_{BL1})+(1-\pi_{BH}-\pi_{BL1})(\pi_{BH}+\gamma\pi_{BL1}) \\
& +[\gamma(\pi_{WH}-1+\pi_{BL1}+\pi_{BH})+\pi_{WL}](\pi_{BL}-\pi_{BL1})
\end{aligned}
$$

Taking the derivative with respect to π_{BL1}

$$
\frac{d}{d\pi_{BL1}}=0=-(\theta_1-\theta_2)-\pi_{BH}+\gamma(1-\pi_{BH}-2\pi_{BL1})+\gamma\pi_{BL}-\gamma(\pi_{WH}-1+2\pi_{BL1}+\pi_{BH})-\pi_{WL}
$$

Combining terms yields

$$
0=-(\theta_1-\theta_2)-(1+2\gamma)\pi_{BH}+2\gamma-4\gamma\pi_{BL1}+\gamma\pi_{BL}-\gamma\pi_{WH}-\pi_{WL}
$$

Solving for the fraction of B types at school 1, $\pi_{BL1}+\pi_{BH}$

$$
\begin{aligned}
\pi_{BL1} &= -\tfrac{(\theta_1-\theta_2)}{4\gamma}+\tfrac{1}{2}-\tfrac{(1+2\gamma)\pi_{BH}}{4\gamma}+\tfrac{\pi_{BL}}{4}-\tfrac{\pi_{WH}}{4}-\tfrac{\pi_{WL}}{4\gamma} \\
\pi_{BL1}+\pi_{BH} &= -\tfrac{(\theta_1-\theta_2)}{4\gamma}+\tfrac{1}{2}-\tfrac{(1-2\gamma)\pi_{BH}}{4\gamma}+\tfrac{\pi_{BL}}{4}-\tfrac{\pi_{WH}}{4}-\tfrac{\pi_{WL}}{4\gamma}
\end{aligned}
$$

We want to show that $\pi_{BL1}+\pi_{BH}$ is less than the inter-racial contact maximizing level at school 1. Let $\pi^*_{BL1}+\pi^*_{BH}$ denote the inter-racial contact maximizing level at school 1 found in Equation 8.

$$\begin{aligned}
\pi_{BL1} + \pi_{BH} &< \pi^*_{BL1} + \pi^*_{BH} \\
-\frac{(\theta_1 - \theta_2)}{4\gamma} + \frac{1}{2} - \frac{(1-2\gamma)\pi_{BH}}{4\gamma} + \frac{\pi_{BL}}{4} - \frac{\pi_{WH}}{4} - \frac{\pi_{WL}}{4\gamma} &< \frac{1}{2} - \frac{(1-2\gamma)\pi_{BH}}{4\gamma} + \frac{\pi_{BL}}{4} - \frac{\pi_{WH}}{4} - \frac{\pi_{WL}}{4\gamma} \\
\underbrace{-\frac{(\theta_1 - \theta_2)}{4\gamma}}_{(-)} &< 0 \qquad \square
\end{aligned}$$

B Tables

Table 11: List of institutions used from the 1989 College & Beyond Database

Duke University
Georgetown University
Kenyon College
Miami University (Ohio)
Notre Dame University
Oberlin College
Princeton University
Pennsylvania State University, State College
Stanford University
University of Michigan, Ann Arbor
University of North Carolina, Chapel Hill
University of Pennsylvania
Vanderbilt University
Washington University, Saint Louis
Wesleyan University
Williams College
Yale University

†See text for details on how the simulations were conducted.

Table 12: Estimates of the Structural Model with Different Values for the Number of Potential Friends[†]

			Specification				
$R^{\ddagger}$	λ_0	λ_1	Intercept	Sex_i	$\lvert A_n - A_i \rvert$	σ	Log likelihood
100	0.0197	0.0737	2.7282	-0.4029	-3.9718	0.3589	4746.4
90	0.0219	0.0818	2.7284	-0.4028	-3.9700	0.3595	4746.4
80	0.0246	0.0918	2.7291	-0.4027	-3.9690	0.3591	4746.4
70	0.0281	0.1045	2.7291	-0.4025	-3.9658	0.3591	4746.4
60	0.0327	0.1214	2.7295	-0.4022	-3.9621	0.3593	4746.4
50	0.0392	0.1448	2.7308	-0.4020	-3.9584	0.3594	4746.4
40	0.0490	0.1794	2.7314	-0.4015	-3.9506	0.3595	4746.4
30	0.0651	0.2356	2.7332	-0.4008	-3.9386	0.3596	4746.3
20	0.0971	0.3427	2.7359	-0.3991	-3.9130	0.3608	4746.2

[†] See text for details on how the simulations were conducted.

[‡] R refers to the number of simulated individuals.

References

[1] Akerlof, G. and R. Kranton (2000) "The Economics of Identity" *Quarterly Journal of Economics*, v. 115, No. 3, pp. 715-753.

[2] Arcidiacono, P. (2005) "Affirmative Action in Higher Education: How do Admission and Aid Rules Affect Future Earnings" *Econometrica*, v. 73, No. 5, pp. 1477-1524.

[3] Arcidiacono, P. and J. Vigdor (forthcoming) "Does the River Spill Over? Estimating the Economics Returns to Attending a Racially Diverse College" *Economic Inquiry*.

[4] Black, D., K. Daniel and J. Smith (2001) "Racial Differences in the Effects of College Quality and Sutdent Body Diversity on Wages" in *Diversity Challenged*, Harvard Educational Review.

[5] Bowen, W.G. and D. Bok (1998) *The Shape of the River: Long-Term Consequences of Considering Race in College and University Admissions.* Princeton University Press.

[6] Carli, L.L., R. Ganley and A. Pierce-Otay (1991) "Similarity and Satisfaction in Roommate Relationships." *Personality and Social Psychology Bulletin*, v. 17 pp. 419-426.

[7] Camargo, B., Stinebrickner, R., and T. Stinebrickner (2009) "Affirmative Action and Interracial Friendships." working paper.

[8] Epple, D., R. Romano, and H. Sieg (2008) "Diversity and Affirmative Action in Higher Education." *Journal of Public Economic Theory* v. 10(4), pp. 475-501.

[9] Foster, J.R. (2005) "Making Friends: A Non-experimental Analysis of Social Group Formation." *Human Relations*, v. 58 pp. 1443-1465.

[10] Hajivassiliou, V.A. and P.A. Ruud (1994) "Estimation Methods Using Simulation," in *Handbook of Econometrics, Vol. 4*, eds. R. Engle and D. McFadden, New York: Elsevier Science.

[11] Henderson, M. and A. Furnham (1985) "Similarity and Attraction: The Relationship between Personality, Beliefs, Skills, Needs, and Friendship Choice." *Journal of Adolescence* v. 5 pp. 111-123.

[12] Kandel, D.B. (1978) "Homophily, Selection, and Socialization in Adolescent Friendships." *American Journal of Sociology* v. 84 pp. 427-436.

[13] Kane, T. (1998) "Racial and Ethnic Preferences in College Admissions." in Christopher Jencks and Meredith Phillips (eds.) *The Black-White Test Score Gap*, Brookings Institution.

[14] Marmaros, D. and B. Sacerdote (2006) "How do Friendships Form?" *Quarterly Journal of Economics*, v. 121, No.1, pp. 79-119.

[15] Mayer, A. and S. Puller (2008) "The Old Boy (and Girl) Network: Social Network Formation on University Campuses." *Journal of Public Economics*, 92 , 329-347.

[16] McPherson, J.M. and L. Smith-Lovin (1988) "Homophily in Voluntary Organizations: Status Distance and the Composition of Face-to-Face Groups." *American Sociological Review* v. 52 pp. 370-379.

[17] Rothstein, J. and A. Yoon (2006) "Mismatch in Law School." working paper.

[18] Weinberg, B. (2005) "Social Interactions with Endogenous Association." working paper.