Assessing the “Mismatch” Hypothesis:
Differentials in College Graduation Rates by Institutional Selectivity*

Sigal Alon
Tel Aviv University

Marta Tienda
Princeton University

Running Header:
Assessing the Mismatch Hypothesis

*We gratefully acknowledge research support from the Mellon Foundation and institutional support from the Office of Population Research at Princeton University. Please direct all correspondence to Sigal Alon, Department of Sociology, Tel-Aviv University, Israel [salon1@post.tau.ac.il]

May 2005
Assessing the “Mismatch” Hypothesis: Differentials in College Graduation Rates by Institutional Selectivity

Abstract: This paper evaluates the “mismatch” hypothesis advocated by opponents of affirmative action, which predicts lower graduation rates for minority students who attend selective postsecondary institutions compared with those who attend colleges and universities where their academic credentials are better matched to the institutional average. Using two nationally representative longitudinal surveys (HS&B and NELS:88) and a unique survey of students enrolled at selective and highly selective institutions (C&B), we test the “mismatch” hypothesis by implementing a robust methodology that jointly considers enrollment in and graduation from selective institutions as interrelated outcomes. Our findings do not support the “mismatch” hypothesis for black and Hispanic (as well as white and Asian) students who attended college during the 1980s and early 1990s.
Assessing the Mismatch Hypothesis

Introduction

Waning support for affirmative action during the 1990s challenged the research community to justify consideration of race as a legitimate factor in college admissions decisions. Proponents of affirmative action practices argue that there is a compelling societal interest in maintaining affirmative action policies because such initiatives are addressing long-term societal needs by promoting diversity and equality of educational opportunity. From this vantage point, race sensitive admissions are not only necessary to correct past discrimination, but also to broaden educational access to disadvantaged students. Moreover, these initiatives enrich the education of all students by assembling a student body with diverse talents and perspectives. Also, responding to a societal need for more members of minority groups in business, government, and in the professions, elite institutions that admit minorities give them opportunities to become leaders in all walks of life (Bowen and Bok, 1998).

Critics of affirmative action charge that racial preferences are contrary to core American values regarding fairness, equality and respect for individual achievement in a meritocracy. They argue that affirmative action constitutes reverse discrimination that lowers the odds of admission for ‘better’ qualified white students. Another claim is that affirmative action stigmatizes all members the target group as unqualified, which results in demoralization and substandard performance regardless of individual qualifications.

One of the main allegations against affirmative action is that treating race and ethnicity as a plus factor for purposes of admission to selective and highly selective institutions sets up minorities for failure because they are putatively unprepared to
succeed academically (Graglia 1993; Sowell 2003; Thernstrom and Thernstrom 1997). Bowen and Bok (1998) and others dubbed this the “fit” hypothesis, but the “Mismatch Hypothesis” better captures the essence of the debate, namely that minority students with lower credentials than the institutional average are “mismatched” at selective institutions and thus have worse outcomes.

The debate about affirmative action practices raises a specter of educational and social policy questions that deserve a careful consideration. As Bowen and Bok (Bowen & Bok 1998: 275) eloquently argue: "In the face of what seems like a veritable torrent of claims and counterclaims, there is much to be said for stepping back and thinking carefully about the implications of the record to date before coming to settled conclusions." This analysis heeds their advice and informs the debate by empirically evaluating the mismatch hypothesis. Using two nationally representative longitudinal surveys (HS&B and NELS:88) and a unique survey of students enrolled at selective and highly selective post-secondary institutions (C&B), we test claims that considering Hispanic origin and race in admissions decisions predestined “affirmative admits” for inevitable failure.

The "Mismatch Hypothesis"

The controversy about affirmative action in college admissions, and the derivative “mismatch” hypothesis, is about access to the most selective postsecondary institutions in the nation. Although many factors determine the selectivity of admissions, two key indicators are singled out for ranking postsecondary institutions, namely the average SAT
score of freshman classes and the percent of applicants admitted (Barron’s 2003; Bowen and Bok 1998; Greenberg 2002; U.S. News and World Report 2003). Based on the 1650 institutions listed in the 2003 Barron’s Guide, only 64 institutions, or 3.9 percent, are classified as “most competitive.” According to Greenberg (2002: 526), the nation’s twenty-five most highly selective universities offer about 50,000 slots annually. Rising demand for a relatively fixed number of slots at the most competitive institutions has certainly fueled growing disapproval of affirmative action in college admissions, but so too has the growing belief that standardized scores on college entrance exams are reliable criteria for establishing admission cutoffs. That black and Hispanic students typically average lower scores on standardized college entrance exams is used to justify claims about their unsuitability to attend selective institutions, where the average SAT score is well above the minority group average (Arkes 1999; Graglia 1993; Lerner and Nagai 2001; Pell 2003).

The basic claim is that the lower average graduation rates of “affirmative admits” result from a mismatch between their academic preparation—indicated by their lower standardized college entrance test scores and high school grades—and the scholastic requirements of the schools that admitted them by taking race into account (Pell 2003). Thernstrom and Thernstrom argue that: "When students are given a preference in admission because of their race or some other extraneous characteristics, it means that they are jumping into a competition for which their academic achievements do not qualify them and many find it hard to keep up." (1997: 406). Presumably, “mismatched” minority students become demoralized, underperform, and ultimately fail to graduate.
Assessing the Mismatch Hypothesis

(Crawford 2000; Thernstrom and Thernstrom 1999; Lerner and Nagai 2001; Pell 2003).

This logic implies that a better match between the academic credentials of minority students with the average of institutions they attend will lead to stronger performance, including higher graduation rates and post-graduate activities (Arkes 1999). In this paper we assess whether minority students are hurt, in terms of graduation likelihood, from attending selective postsecondary institutions compared to their same-race counterparts who attend less selective colleges and universities where their academic credentials are presumably better matched to the institutional average.

The claims embedded in the “mismatch” hypothesis contrast with both common knowledge and empirical research based on elementary and secondary school students, which demonstrates that, regardless of prior achievements, students who attend higher tracks and/or better schools make greater scholastic gains (Gamoran 1987; Gamoran and Berends 1987; Hallinan 1996; 2001; Entwisle et. al. 1997; Hoffer, 1992; Gamoran and Mare, 1989). Empirical studies have demonstrated repeatedly the advantages of placement in higher ability groups with better instruction, less distraction, more time spent on task, more academic role models, and more serious learning climates (Hallinan 2001). These findings indicate that cognitive skill development depends crucially on the opportunities for learning that schools afford (Gamoran 1987). Participation in higher academic tracks and more demanding schools may have offsetting advantages for disadvantaged students (Hallinan 2001).

Black students’ postsecondary experiences further support this idea. By demonstrating that for all intervals of the SAT distribution graduation rates of black
Assessing the Mismatch Hypothesis

students increase as institutional selectivity rises, Bowen and Bok (1998) challenged the core of the “mismatch” hypothesis. Not only do their findings dispute allegations that black students can not succeed at selective colleges and universities, but they also demonstrate a consistent positive association between institutional selectivity and several post-graduation outcomes, including completion of advanced degrees, earnings, and overall satisfaction with college experiences (Dworkin 1998). Attending to the same question, Kane (1998) argues that affirmative action narrows rather than widens gaps in college retention rates by race because the net relationship between college selectivity and college graduation rates is positive for all students.

While informative, findings based on comparisons of students attending institutions that differ in the selectivity of their admissions fall short of demonstrating a causal link between institutional selectivity and subsequent college outcomes by ignoring student allocation to selective and nonselective college destinations (Winship and Morgan, 1999). Because the determinants of attending a selective institution and graduating from college overlap to a large extent, ignoring students’ institutional assignment can lead to biased inferences about the effect of college selectivity on graduation resulting from unmeasured influences that are correlated both with institutional assignment and college graduation. This is particularly important if the conditional probability of attending selective postsecondary institutions differs for minority and nonminority populations of varying academic qualifications—that is, if institutions practice affirmative action. Therefore, testing the mismatch hypothesis requires joint estimation of the selection regime allocating students to colleges and their
Assessing the Mismatch Hypothesis

graduation probability. Jointly modeling the graduation outcomes and the institutional assignment process not only resolves a methodological weakness of prior studies, but bears important substantive implications for the ongoing debate about affirmative action.

Our evaluation of the “mismatch” hypothesis extends prior research in three important ways. First, unlike prior studies, we account for the selection bias stemming from the fact that the allocation of individuals to colleges and universities that differ in the selectivity of their admissions is endogenous to subsequent academic success. Our analytical framework—devised to assess the causal link between institutional selectivity and college graduation—stands to this challenge by combining several statistical methods that model the allocation regime governing assignment to a selective institution (Heckman and Hotz, 1989). We use two different estimation techniques, namely Propensity Score (Rosenbaum and Rubin, 1983) and Matching Estimator (Heckman et al 1997; 1998) that differ in their assumptions about functional form. We also present evidence that takes into account the possibility of selection on unobservables. Using a dummy endogenous variable model (Heckman, 1978), we simultaneously model enrollment in selective and nonselective institutions on the likelihood of graduation.

Second, we compare the experiences of students attending selective and highly selective institutions using the College and Beyond (C&B) database to national college-bound populations. Analyzing the C&B data allows us to analyze the experiences of students attending the most competitive institutions in the country, which are the main focus of the controversy about race-sensitive admissions practices. Comparing the experiences of those students to the national college-bound populations yields a broader
Assessing the Mismatch Hypothesis

perspective on the selection processes that allocate students to institutions that vary in the competitiveness of their admissions, and also provides a reliability check on Bowen and Bok’s (1998) claims for blacks based on the C&B data. This is especially important because the C&B institutional selectivity spectrum is truncated, which may limit the generalizability of inferences.

Finally, we compare the educational experiences of the four major ethno-racial groups. By emphasizing black-white differences, prior analysts neglect the most rapidly growing segments of the college population, namely Hispanics and Asians (see Bowen and Bok, 1998; Kane, 1998). We pay special attention to Hispanics, whose position in higher education has been relatively understudied compared with blacks. Although Hispanics are notorious for their educational underachievement, particularly their persistent and elevated rate of high school non-completion (Fry 2003), understanding their success in higher education—both who gains admission to and who graduates from selective institutions—holds promise for the design of policy geared to improve their standing in postsecondary institutions and beyond. Multi-group comparisons based on nationally representative data not only situate Hispanics in the broader terrain of higher education, but also validate findings for blacks based on a subset of highly selective institutions.

In what follows, we elaborate on the testable implications of the “mismatch” hypothesis and formulate a strategy for evaluating them. After describing the three data sources used for empirical estimation, we report the statistical results. We find that conditional on admission, all groups of students who attend selective institutions are more
likely to graduate within six years of enrollment than their counterparts who attended less selective colleges. We reject the “mismatch” hypothesis for students who enrolled at the most selective institutions during the late 1980s and early 1990s.

**Assessing the "Mismatch Hypothesis"**

The broad implication of the mismatch argument is that Hispanics and blacks enrolled in institutions with selective and highly selective admissions have a lower probability of graduating than their counterparts with similar characteristics who attend post secondary institutions with less selective admission criteria. Before developing our analytical strategy, it is important to distinguish between racial and ethnic graduation probabilities within institutions and the predictions of the “mismatch” hypothesis, which focus on same group comparisons between selective and nonselective institutions. Both sources of inequality are of policy interest, but we are concerned with the latter.

Assessing the “mismatch” hypothesis requires comparing the graduation likelihood of students attending selective institutions with their same-race counterparts who attend nonselective schools. However, graduation from a selective institution is conditional on being admitted, a highly selective process influenced by many observed and unobserved factors that also influence the likelihood of timely college graduation. Analytically, the mismatch hypothesis has two related, but separable components: (1) group differences in the probability of graduating from institutions that differ in the selectivity of their admissions; and (2) group differences in the probability of attending a selective institution. The following two equations formally summarize the model.
Assessing the Mismatch Hypothesis

\[
\Pr( Y_i = 1 | X = x_i ) = \alpha D_i + \beta' X_i + \varepsilon_i \tag{1}
\]

\[
\Pr( D_i^* = 1 | Z = z_i ) = \phi' Z_i + \nu_i \tag{2}
\]

Let \( Y \) be a measure of college graduation of the \( i \)th individual; \( D_i ^* \) is a latent continuous variable denoting institutional selectivity of the \( i \)th individual's destination college, where \( D_i = 1 \) if \( D_i ^* \geq 0 \) (selective college destination) and \( D_i = 0 \) if \( D_i ^* < 0 \) (nonselective college destination); \( \alpha \) is its coefficient; \( X \) is a vector of observed attributes that influence college graduation; \( \beta \) is the vector of their coefficients; and \( \varepsilon \) is an error term that captures unobserved factors affecting graduation.

The parameter \( \alpha \) in equation (1) is the coefficient of interest for testing the "mismatch" hypothesis: the hypothesis cannot be rejected when \( \alpha < 0 \); that is, if students attending less selective schools (\( D = 0 \)) are more likely to graduate compared with students attending more demanding institutions (\( D = 1 \)). However, comparisons between students who attend colleges that differ in selectivity are prone to bias because of observable and unobservable differences that govern the selection regime. Equation (2) represents the selection regime that determines assignment into college destinations, where \( Z \) is a vector of observed exogenous covariates that allocate students among postsecondary institutions; \( \phi \) is a vector of associated coefficients, and \( \nu \) is an error term that captures unobserved influences in the allocation regime.

We use the Counterfactual Account of Causality conceptual and notational framework to describe the problem (for a review see Winship and Morgan, 1999; Winship and Sobel, 2004; Heckman et al., 1997). Each individual is exposed to one of
two destination states, i.e. nonselective (D=0) vs. selective colleges (D=1), although each individual may be exposed to either state a priori. Each college destination is characterized by a set of conditions that, upon exposure, affects students’ likelihood of college graduation. In this framework, college destination states may be considered as treatment (D=1) and control (D=0) in a quasi-experiment. A key assumption of the counterfactual framework is that all students have a graduation probability in both states: the one in which they are observed and the one in which they are not observed. A student attending a nonselective college has an observable outcome in a nonselective college, $Y_i^0$, and an unobservable counterfactual outcome at a selective institution, $Y_i^1$.

The gain from attending a selective institution is given by $\delta_i = Y_i^1 - Y_i^0$. If $Y_1$ and $Y_0$ could be observed simultaneously for all students, there would be no evaluation problem because $\delta$ would be measured for everyone. The evaluation problem arises because ordinary observational data do not provide sample counterparts for the missing counterfactual, namely $Y_0$ values for students who attend a selective institution (D=1). Because only one outcome is observed for each individual, it is not possible to calculate individual-level gains from attending a selective institution. The evaluation problem is therefore a missing data problem (Heckman et al., 1997).

However, we can calculate the mean outcome for all students who attend selective and nonselective institutions, $\bar{Y}^1$ and $\bar{Y}^0$, respectively. The standard estimator for the average treatment effect is the difference between these two estimated means, $\bar{\delta} = \bar{Y}^1 - \bar{Y}^0$. This strategy is adequate for randomized studies in which the random
assignment insures that the treatment assignment is independent of all the individual characteristics. In other words, a balanced design – a design where the treatment and control groups do not differ with respect to Z in any way that is also related to the treatment assignment D – is required to estimate $\delta$. However, because students allocated to selective and nonselective institutions are dissimilar in their observed characteristics, such a straightforward comparison is not possible (Slavin, 1990). In this case, some observed set of factors in $Z$ is related to the likelihood of graduation. Therefore, in observational studies we must adjust for possible differences in observed Z attributes across treatment and control groups to ensure that the assignment is random with respect to the graduation outcome. The ignorability condition is sufficient for the standard estimator to be unbiased and consistent (Rosenbaum and Rubin, 1983).

We use two methods that adjust for observed group differences in Z. Because there is no perfect model to assess causality in nonexperimental designs, and because alternative methods may yield different results, our multi-pronged analytical strategy that builds on different assumptions about underlying relationships affords a highly rigorous standard for testing the mismatch hypothesis (Winship and Mare, 1992; Heckman and Hotz, 1989). The continuing level of controversy requires nothing less.

The first approach is the widely used matching estimator (Heckman et al., 1997; 1998). To match, one identifies individuals in the treatment and control groups with equivalent or at least similar values of the $Z$ covariates and matches them, creating a new sample of matched cases. The standard estimator is then applied to the matched sample. By construction, the treatment and control cases in the matched sample have identical
values of $\bar{Z}$. Thus, matching estimates the effect of any potential differences in the
distribution of $\bar{Z}$ between the treatment and control groups by equating the distribution
of $Z$ across students who enroll in selective and nonselective institutions (Winship and
Sobel, 2004). Calculating $\delta$ for students matched on $\bar{Z}$ simulates a balanced design in
which the assignment is random with respect to the graduation outcome.

Matching has several advantages. First, it is a non-parametric estimator that
makes no assumptions about the functional form of the dependence between the outcome
of interest and $\bar{Z}$. Second, matching insures that the Zs in the treatment group are similar
(matched) to those in the control group. Third, because fewer parameters are estimated
than in a regression model, matching is more efficient. One problem with the traditional
matching approach is that, if the sample is small the there are more than few covariates in
$\bar{Z}$, it may be difficult to find both treatment and control cases that match.

A solution is suggested by Rosenbaum and Rubin (1983) who advocate the
Propensity Score method. Propensity scores represent the probabilities that individuals
with observed characteristics $\bar{Z}_i$, are assigned to selective rather than nonselective
institutions. They argue that little is gained by refined matching on the $\bar{Z}$ variables, as in
matching estimations compared with the propensity scores alone. Propensity scores
contain all the information needed to create a balanced design, while matching can be
estimated on a single dimension (i.e. the propensity score), which makes the matching
more feasible. If the likelihood of attending a selective college is purely a function of the
observables, then conditional on the $\bar{Z}$ vector, assignment is random with respect to the
Assessing the Mismatch Hypothesis

graduation outcome. Including the propensity score as a control variable in the graduation equation (equation 1) removes from both Y and D the component of their correlation that is due to the assignment process (Winship & Morgan 1999). By so doing, the parameter α in equation (1) gives us \( \delta \).

Testing the “mismatch” hypothesis once selection is accounted for requires evidence of lower graduation odds for students attending selective schools relative to statistically similar group members attending less demanding schools. That is,

\[
H_0 : \quad \delta < 0 \mid Z = z_i \tag{3}
\]

\[
H_1 : \quad \delta \geq 0 \mid Z = z_i \tag{4}
\]

The “mismatch” hypothesis \( (H_0) \) predicts that Hispanic and black students (as well as whites and Asians) are less likely to graduate from selective compared with nonselective institutions, conditional on their likelihood of attending a selective institution. Because the predictions of the “mismatch” hypothesis focus on same group comparisons attending selective and nonselective institutions, estimation of \( \delta \) is group-specific. The “mismatch” hypothesis cannot be rejected if group counterpart students are more likely to graduate from less selective schools \( (\delta < 0) \). Alternatively, we can reject the “mismatch” hypothesis \( (H_1) \) if statistically similar Hispanic and black students are equally or more likely to graduate from selective compared with nonselective institutions (when \( \delta \geq 0 \)).

Both propensity scores and the matching estimator are designed to address selection on observable characteristics \( Z \). However, selection could also be related to unobservable attributes. That is, when unobserved (by the analyst) variables influence
both the institutional assignment and the graduation outcome, the errors of the prediction equations for both outcomes ($\varepsilon$ and $\nu$) are likely to be correlated, and hence must be jointly estimated to obtain unbiased estimates (Pindyck and Rubinfeld 1998). We examine this issue of selection on unobservables to get a sense about (1) how much such a selection is a problem in our context; and (2) how much and to what direction estimates of $\delta$, based on propensity score and the matching estimator would change after such a correction. Above all we want to know whether these estimates are upwardly or downwardly biased.

Ensuring the ignorability condition for unobservable attributes is more complicated compared with observed characteristics. None of the existing alternatives is perfect. To assess the bias stemming from the "selection on the unobservables" ($\varepsilon$ and $\nu$ are correlated), we jointly estimate the likelihood of enrollment at selective institutions (equation 2) and the likelihood of graduation (equation 1). Following Heckman (1978), we fit a dummy endogenous variable model to obtain unbiased and consistent estimates. Using appropriate distributional assumptions this technique estimates error covariances (i.e. the correlation between $\varepsilon$ and $\nu$) across equations (Pindyck and Rubinfeld, 1998; Greene, 2000). There are some exclusion restrictions, but since the determinants of college enrollment and graduation overlap to a large extent, identification is achieved primarily via distributional assumptions (Greene, 2000).
Assessing the Mismatch Hypothesis

Data and Empirical Estimation

To estimate graduation probabilities for students attending selective and highly selective institutions, we analyze the 1989 cohort of the College and Beyond (C&B) database. However, because these data represent the experience of a relatively small share of students attending selective and highly selective four-year institutions, we also analyze two nationally representative data sets—the High School and Beyond (HS&B) and the National Educational Longitudinal Survey (NELS:88).

C&B is a restricted-access database constructed by the Andrew W. Mellon Foundation between 1995 and 1997 (Bowen and Bok 1998: Appendix A). Two strengths of these data for analyzing graduation rates at selective colleges and universities are the accurate persistence data derived from college transcripts (rather than students’ self-reports) and the relatively large samples of minority students attending highly selective institutions. The core of the C&B database is an “institutional data file” consisting of undergraduate students who enrolled at one of 28 academically selective colleges and universities in the fall of 1989.7

The institutional file contains information drawn from students’ applications and transcripts, including race, sex, SAT scores, college grade point average, major field of study, and importantly, graduation status (outcome and date). Institutional records were collected for all students who enrolled in the fall of 1989 at all but three of the C&B institutions.8 Individual student records are linked to several other sources, including a survey that collected retrospective data, files provided by the College Entrance Examination Board (CEEB), and data collected by Higher Education Research Institute...
Assessing the Mismatch Hypothesis (HERI) at the UCLA. Limiting the analysis to U.S. residents/citizens with valid racial and ethnic identities and graduation status, the final sample of 29,018 students includes 23,086 white, 2,260 black, 1,235 Hispanic, and 2,437 Asian origin students.

The HS&B and NELS:88 surveys are nationally representative samples of the 1982 and 1992 high school graduation cohorts, respectively, collected by the National Center for Educational Statistics (NCES). The detailed education histories provided by these longitudinal surveys make them ideal for studying both the transition to college and the institutional selectivity of matriculants. In addition to over-samples of blacks and Hispanics, these surveys include rich information about test scores and academic high school performance, as well as standard indicators of family background. Transcript data are available for students who attended college.

Our substantive interest dictates restricting both samples to students who attended a 4-year postsecondary institution with a valid institutional selectivity ranking. For the NELS survey, the analysis sample includes the 1992 high school graduation cohort respondents who were interviewed in the 2000 follow-up. Of the 4,530 eligible students, 3,326 are white, 386 black, 361 Hispanic, and 457 of Asian origin. The comparable HS&B analysis sample consists of 4,704 students who graduated from high school in 1982 and who were re-interviewed 10 years later. Eligible respondents include 3,260 whites, 644 blacks, 559 Hispanics, and 241 Asians.

Descriptive analyses are weighted to adjust for over-sampling, non-response, and attrition. Moreover, multivariate analyses are adjusted to account for the clustered survey design of the data set. Flags for missing values are included in all models, but are not
reported in the results presented here. Appendix A provides detailed definitions and descriptive statistics of variables analyzed from each dataset.

Variables: The measure of institutional selectivity for the HS&B and NELS:88 is identified as the mean combined SAT score of entering freshmen in 1982 and 1992, based on the Cooperative Institutional Research Project (CIRP) data. "Selective" institutions include those whose mean class SAT scores were above 1050; the remainder of 4-year institutions are classified as "Nonselective." The mean combined SAT score of the 1989 entering freshmen exceeded 1050 for all C&B institutions, which qualify as selective under the CIRP classification. The C&B data defined "Most selective" institutions as those where the average combined SAT score of the entering class was 1300 or higher (Bowen and Bok, 1989, Appendix B).

The graduation equation (equation 1) includes a dummy variable for attending a selective school and a vector of covariates that are known to influence college persistence and success (Tinto, 1993); including: social class (parental education and income); academic preparation (high school class rank, individual SAT scores) to isolate student qualifications from institutional selectivity; and sex. The propensity score model (equation 2) includes Z covariates known to influence access to selective institutions. These covariates are also used to match students in the matching estimator procedure. These covariates are: background characteristics (parental education and income, dummy variables for public high school attendance, home geographic region, a dummy for residing in a rural area, athlete status), and academic preparation based on high school class rank and SAT scores (Persell et al. 1992; Davies and Guppy 1997; Hearn 1984;
Assessing the Mismatch Hypothesis

1990; 1991; Karen 2002; Kingston and Lewis 1990; Carnevale and Rose 2003). Race and ethnicity dummy variables are included only in pooled models to capture minority students’ possible preferential admission advantage.

Results

Table 1 reveals the diversification of college campuses during the post-Bakke expansion of affirmative action. The relative share of white students attending selective institutions declined since the early 1980s, when the HS&B students began their postsecondary education. Asians witnessed the most substantial gains at selective institutions, nearly doubling their shares between 1982 and 1992. Despite the rapid growth of the Hispanic college-age population since 1980 and the aggressive recruitment of black and Hispanic students by admissions officers from selective colleges and universities, their shares of the entering cohort rose only slightly during the 1980s.

[Table 1 about here]

At the most competitive institutions included in the C&B sample, the combined representation of blacks and Hispanics reached 13 percent. Paralleling national trends, the diversification of the most elite institutions largely reflects the rising Asian presence. Cross-data comparisons reveal relatively similar shares of black students attending selective institutions, approximately 6 to 7 percent for NELS and C&B, respectively, but lower shares of Asian and Hispanic enrollees in selective institutions included in the C&B data compared with NELS. Most likely this reflects the exclusion of Texas and California public flagship institutions from the C&B database.
Although minority enrollment and graduation from selective postsecondary institutions increased since 1980, racial and ethnic disparities in graduation rates persist. Six-year graduation rates are higher, on average, at selective compared with nonselective institutions, which undermines allegations that lowering admission thresholds to include more minority students lowers overall graduation rates. Table 2 refutes this claim by showing that black and Hispanic student graduation rates increased between 1982 and 1992 at both selective and nonselective institutions, even as their relative shares of the student body rose. Specifically, the graduation rate of black students attending nonselective schools rose from 26 percent in 1982 to 48 percent in 1992, while black graduation rates at selective institutions rose 20 percentage points—from 52 to 72 percent over the decade. The rise in Hispanic graduation rates during this period rose from 26 to 40 percentage points at nonselective institutions compared with a modest increase of only 7 percentage points at the selective institutions.¹¹

Graduation rates are uniformly higher for selective compared with nonselective institutions (see “odds ratio” column), although the graduation gap between selective and nonselective colleges and universities narrowed over time for all groups. White students’ 1982 graduation odds ratio of 3.9 indicates that students attending selective schools were about 4 times more likely to graduate than their race counterparts attending nonselective institutions. This odds ratio declined modestly ten years later. For Hispanic and black students the 1982 graduation odds ratios between selective and nonselective institutions were somewhat smaller than for whites (2.7 and 2.9, respectively) and remain stable over
time—increasing slightly for blacks and decreasing slightly for Hispanics. That graduation rates are higher for the C&B institutions than the selective schools included in the NELS and HS&B samples is unsurprising given the positive association between institutional selectivity and graduation rates overall (Kane, 1998). However, these differences also highlight a limitation of the C&B data for an exclusive test of the “mismatch” hypothesis; namely, the truncated distribution of institutional selectivity greatly restricts variation in both student attributes and graduation rates.

In sum, Hispanic and black students’ graduation probabilities are higher at selective compared with nonselective institutions, contrary to the predictions of the “mismatch” hypothesis. In fact, it appears that the racial and ethnic graduation gap narrows as institutional selectivity increases. However, this conclusion is tentative because students attending selective institutions are generally better prepared academically than their counterparts enrolled in nonselective institutions. Thus the higher minority graduation rates at selective institutions may simply reflect their higher average qualifications relative to their group average.

The Mismatch Regime

To assess the degree of mismatch between students and their postsecondary institutions, Table 3 reports the deviation of each group SAT and class rank mean from the institutional tier average. The top panel reports the corresponding tier averages to which group-specific means are compared. In both 1982 and 1992, average SAT scores for white and Asian students were slightly higher than the institutional tier average at
Assessing the Mismatch Hypothesis

...both selective and nonselective institutions. By contrast, and in line with the claims of mismatch proponents, both black and Hispanic students enrolled in selective institutions averaged test scores well below the respective institutional tier averages—162 and 112 points lower, respectively, for black and Hispanic HS&B respondents, and 176 and 95 points lower, respectively, for NELS students. However, the deviation of the black and Hispanic mean SAT scores for the nonselective institutional average was greater still, about 180 and 115 points for blacks and Hispanic HS&B students, respectively, and 155 and 100 points for blacks and Hispanic NELS students, respectively. Racial and ethnic disparities in class rank mirror those based on average SAT scores.

[Table 3 About Here]

That black and Hispanic mean SAT scores lag behind institutional averages at both selective and nonselective institutions would appear to challenge the “mismatch” hypothesis. Presumably, group-specific disparities in test scores lower the odds that black and Hispanic students will graduate in six years, yet the tabular differences reveal more pronounced disparities among students attending nonselective compared with those attending selective institutions. The C&B results show similar patterns, indicating that Hispanic students are better matched to their institutions’ average academic level than are blacks, but less well than either whites or Asians. Hispanic students’ mean SAT scores lag about 83 and 89 points, respectively, behind the tier averages at selective and most selective schools, but they are better matched to most selective institutions than to selective institutions based on class rank. That Hispanic and black students attending both selective and nonselective institutions appear to be less well prepared scholastically, has
direct implications for their persistence in the postsecondary education system and ultimate graduation. Their unequal preparation may be more consequential at the most selective institutions compared with less selective schools, according to the “mismatch” hypothesis, because the academic curriculum is more demanding (Bowen and Bok 1998; Massey, et al. 2003).

Comparisons between academic preparedness of students attending selective and nonselective institutions underscores the nature of misunderstanding about the “mismatch” hypothesis. Although minority students average lower test scores and class rank than whites and Asians, those enrolled in selective colleges are better prepared than their same-race counterparts who attend nonselective institutions. The “mismatch” hypothesis is not about racial and ethnic differences in graduation within institutions, but rather about same group comparisons across institutions that differ in the selectivity of their admissions. The ultimate question deriving from the “mismatch” hypothesis is whether Hispanics and blacks enrolled in institutions with selective and highly selective admissions have a lower probability of graduating than their counterparts with similar characteristics who attend institutions with less selective admission criteria. The multivariate results provide more robust evidence about whether race-sensitive admissions harm their beneficiaries than the tabular results.

**Multivariate Analysis**

The multivariate analysis is designed to assess the effect of institutional selectivity on six-year graduation status (1 = yes, 0 = no) of white, black, Hispanic, and Asian
Assessing the Mismatch Hypothesis

students for all three surveys. Because the predictions of the “mismatch” hypothesis focus on same group comparisons attending selective and nonselective institutions, estimation of $\delta$ is group-specific (except for one general pooled model). The strategy is designed to account for selection bias stemming from the fact that the institutional selectivity allocation regime is endogenous to subsequent academic success. To that end, we present results from a propensity score analysis and from a matching estimator.

For the propensity score analysis we calculated a propensity score, as a function of several Z covariates that are known to influence college access: parental education; family income; SAT scores; high school class rank; type of high school; geographic region and a dummy for residing in rural area. We add the calculated propensity score as a control variable in the graduation equation (equation 1), effectively "subtracting out" of $Y_i$ and $T_i$ the component of their correlation due to the assignment process (Winship & Morgan 1999; Rosenbaum and Rubin 1983). To estimate the matching estimator we match individuals on the same set of Z covariates to obtain the average treatment effect.\(^{13}\) For ease of interpretation, Table 4 reports marginal effects associated with the covariate of theoretical interest, namely $\delta$, estimated by both propensity score methods (PS) and matching estimator (ME).\(^{14}\) Also reported in the table are the baseline predicted probabilities of graduating from a nonselective institution, controlling for family background and academic achievements.

Results for the entire HS&B cohort indicate that, regardless of the statistical method employed, there is a positive and substantial effect of institutional selectivity on 6-year college graduation status. That both methods yield very similar results attests that
our results are robust to functional form or other method-specific assumptions. For the entire cohort, the predicted probability of graduating from a nonselective institution is 0.51. The gain from attending a selective institution is substantial as the graduation probability is about 0.18 higher for students who are match on the $Z$ covariates.

Nonetheless, the group-specific models reveal race and ethnic variation in the benefits associated with attending a selective institution in 1982. The estimates of the matching estimator ranges from 0.12 for Hispanics to 0.36 for Asians, but whites and blacks share similar gains in graduation probability, on the order of 0.18-0.19. However, these gains need to be judged against the group-specific baseline probability. Because black and Hispanic students' baseline graduation probabilities are much lower than whites, their gains from attending a selective institution are much more substantial than are those of whites'.

[Table 4 About Here]

Point estimates for the 1992 entering class (NELS:88 data) reveal the same positive relationship between institutional selectivity and college graduation, regardless of estimation method used. For some groups the marginal effects are smaller compared with the 1982 cohort. This could reflect a ceiling effect on the likelihood function that is associated with the rise in the baseline graduation probabilities. Consistent with this interpretation, the marginal effects are smaller for groups with the highest average graduation probabilities, namely whites and Asians. For the entire cohort, the gain in the graduation probability from attending a selective institution is about 0.12. For white students both methods yield the same point estimate (0.10). However, for minority
groups it appears that the functional form of the propensity score method underestimates their gains from attending a selective institution. The matching estimator results suggest that all minority groups benefit more than whites from attending a selective institution in 1992. Hispanics' gains are especially large considering their lower baseline graduation probabilities. Combined, these results in Table 4 allow us to safely reject the "mismatch" hypothesis for all HS&B and NELS:88 students.

To further test the “mismatch” hypothesis, we direct attention to highly selective colleges. Not only are these institutions the focus of controversy about race-sensitive admission policies, but elite institutions also are wealthier and have other resources to support all students they admit, thereby increasing their odds of graduation. For this reason the difference in graduation probabilities between C&B students who attend elite schools and those attending very and highly selective college and universities is particularly instructive. Again, we find a positive causal impact of institutional selectivity on the likelihood of graduation, regardless of the estimation strategy used. The smaller gains compared with the national HS&B and NELS:88 cohorts are to be expected because of a ceiling in the graduation likelihood function and because of the smaller variation in graduation rates among the C&B institutions.

Nevertheless, findings reported in Table 4 reject the mismatch hypothesis for the C&B students. Results based on the non-parametric matching estimator are particularly revealing. Evidence based on comparisons of “matched” samples that attend selective and highly selective institutions implies that black and Hispanics gain more from attending a most selective institution relative to a less selective college compared to whites. For
Assessing the Mismatch Hypothesis

example, the matching estimator point estimate for whites and Asians is 0.07 and 0.04, respectively, whereas it is 0.09 for blacks and 0.11 for Hispanics. This suggests that underrepresented minority students benefit more than whites from attending the most selective schools in the country. Thus, the C&B findings not only refute the “mismatch” hypothesis for all students, but also indicate special gains for minority students from attending elite institutions.

Taken together, evidence from three datasets using two different estimation procedures provides a robust assessment of the causal effect on graduation associated with attending selective institutions, conditional on admission and enrollment. The results clearly demonstrate that attending a selective or highly selective college increases the likelihood of timely graduation, net of initial differences. The results lend no support for the "mismatch" hypothesis for minority students who attend selective and highly selective institutions.

Selection on Unobservables

Admission offices consider many factors, such as legacy status, special talents (e.g. musical ability), and unique circumstances (e.g. hardships or special awards), that are unmeasured by surveys. Therefore, we also consider whether selection on unobservables determines access to different college destinations, and if so, whether the estimates obtained from propensity score and matching estimator techniques are upwardly or downwardly biased. To evaluate possible bias from the "selection on the unobservables" we jointly estimate the likelihood of enrollment at selective institutions
Assessing the Mismatch Hypothesis

(equation 2) and the likelihood of graduation (equation 1). Specifically, we use a bivariate probit technique that fits a dummy endogenous variable model, as suggested by Heckman (1978), to correct for correlated errors across equations (Pindyck and Rubinfeld, 1998; Greene, 2000). The estimated $\rho$ (Rho coefficient) is of special importance for our purpose because it reveals the correlation between the error terms $\varepsilon$ and $v$ of equations (1) and (2), and indicates whether selection on unobservable is a problem, and, consequently, whether the joint estimation is required. Whenever, $\rho$ is not statistically significant, selection on unobservables is assumed to be ignorable. This implies that estimates based on propensity scores and matching estimator are not biased due to different distributions of unobservables.

Table 4 also reports the results from fitting a bivariate probit model to the three datasets. We report the marginal effects associated with attending a selective institution only when the $\rho$ coefficient was found to be statistically significant. Overall, for both the 1982 and 1992 cohorts, the $\rho$ coefficient is not significantly different from zero. That the $\rho$ coefficient is not statistically significant for minority groups (as well as for whites in 1992) suggests that selection on unobservables is not a major underlying factor in the assignment process. $^{15}$ The $\rho$ coefficient is significant only for whites in the HS&B data, indicating selection on unobservables and supporting need for a simultaneous assessment for this group. Specifically, after controlling for the selection on unobservables, the effect of attending a selective institution increases white students’ graduation probability by approximately 40 percent. This suggests that previous estimates
Assessing the Mismatch Hypothesis

based on propensity scores and the matching estimator underestimate whites' benefits of attending a selective institution in 1982.

Results for the C&B data are somewhat different. The $\rho$ coefficient is significant for all minority groups, indicating that selection on unobservables might be involved. The marginal effects produced by the dummy endogenous variable model are larger than those produced by both propensity score and the matching estimator techniques. In other words, taking the selection on unobservables into account, the minority students' gains from attending the most selective institution in the country are even larger than those produced by the other methods.

These results raise questions about why unobserved (by the analyst) characteristics are more influential determinants of graduation in the C&B than in the national data. One possible explanation is the "full file review" in admission performed by elite institutions compared to a more formula-based admission in less selective institutions—particularly before race sensitive admissions practices came under assault in the mid-1990s. Clarke and Shore (2001) argue that selective, private institutions, which base their decision on full file review, tend to interpret a student's test score and diversity characteristics subjectively, and in combination with other information. In this process, each student application is read in its entirely by one or more of the admission staff, who assign a score based on his/her subjective evaluation of all the information in the file. Selective public institutions, that had much larger applicant pools but proportionately smaller admission staff than elite institutions, tend to rely on objective formulas that draw heavily on student's high school achievements and test scores.
Results reported in Table 4 indicate that the most selective institutions pay more attention to a diverse array of (unobserved by us) characteristics when deciding who to admit. The propensity score and matching estimator methods that adjust for observed characteristic are successful in simulating the admission decision in most postsecondary institutions, selective and nonselective. However, these methods fall short of replicating the subjective admission decisions that take place at the most selective institutions in the country. Moreover, Clarke and Shore (2001) note that the full file review is viewed by admissions directors at private selective colleges as the best way to predict whether a student would be successful in college. By showing that, after controlling for unobservables, the benefits from attending an elite institution are larger than simply controlling for observed characteristics, our results lend support to this view.

Conclusion

By rejecting the mismatch hypothesis, our results are consistent with claims that minority students thrive at selective post-secondary institutions, despite their disadvantaged starting lines (Bowen and Bok 1998; Massey et al 2003). Minority students’ graduation likelihood increases as the selectivity of the institution attended rises. Our findings, based on three different datasets and several analytical methods, suggest that the mismatch hypothesis is empirically groundless for black and Hispanic (as well as for white and Asian) students who attended college during the 1980s and early 1990s. Based on the robust evidence presented, we conclude that affirmative action
practices both broaden educational opportunity for minority students, and enable them to realize their full potential.

These findings clearly demonstrate the advantages associated with attending a more selective institution and call for future research to identify the mechanisms that produce such an advantage. Institutional selectivity might reflect better learning opportunities via better prepared classmates or better teachers (Kane 1998). It also could capture student retention possible because of the large institutional endowments that permit smaller classes and facilitate strong mentoring at the most competitive schools. In other words, if the higher graduation rates reflect better opportunities for learning, then it is incumbent on future researchers to specify what "best practices" enable minority students to succeed at less selective institutions.

Support mechanisms are especially important for students from socioeconomically and academically disadvantaged backgrounds, particularly first generation college goers, among whom black, but especially Hispanic students are disproportionately represented (Tinto, 1993). Evidence about interaction with faculty members unequivocally suggests that more contact with professors and others on campus is conducive to higher graduation rates (Pascarella and Terenzini, 1980; Pascarella, et al., 1978; Nettles, 1991; Nettles et. al 1986; Davis, 1991; Von Destinon, 1988; NCES, 1996). Alon (2004) directly links financial aid to college graduation by showing that grants, more than loans, help equalize black and Hispanic minority students’ college success with that of whites and Asians.

Nevertheless, our results do not speak to the persistent racial and ethnic
Assessing the Mismatch Hypothesis

graduation gap within selectivity tiers. Given the immense efforts and ample resources devoted to attracting and recruiting under-represented minorities to the most selective colleges and universities, evidence that any blacks and Hispanic students leave these institutions without a college diploma is disconcerting. Even more disturbing are race and ethnic differences in graduation rates among students of comparable academic ability and socioeconomic background (Bowen and Bok 1998; Small and Winship 2002; Vars and Bowen 1998).

Racial and ethnic gaps in college graduation rates are of major concern not only because education serves as a gateway to personal financial success and social standing, but also because of the shadow that graduation disparities casts on race-sensitive admission practices. For these reasons, researchers must continue to explore reasons for minority students' underperformance in both selective and nonselective institutions. Striving to increase college access while narrowing graduation gaps is all the more urgent in light of the changing demographic contours of the college-age population.
Assessing the Mismatch Hypothesis

References


Assessing the Mismatch Hypothesis


Assessing the Mismatch Hypothesis.


Notes

1 This is manifested by the small industry that developed over the past two decades to prepare students, mainly from middle and upper middle classes, to improve their SAT scores (McDonough, 1994).

2 Bowen and Bok were aware of this limitation and were directly responsible for ensuring that the college-going behavior of Hispanics was studied with the C&B data. This research derives from that effort.

3 A narrower interpretation of the mismatch hypothesis is that the larger the gap between students’ credentials and institution-specific SAT averages, the greater the likelihood that they will not graduate from college. There are several issues that such a narrow interpretation raises (e.g. levels of mismatch, multidimensionality of mismatch, and selection issues). Weighting up the two directions we believe that the broader implications are much more important and better inform the post Grutter policy debate. Our analytical strategy is therefore designed to deal with the broader implications of the mismatch hypothesis.

4 This is one transition in a multi-stage selection process that includes high school graduation and the decision to continue education after high school. It is necessary to control for prior selection stages so as not to confound current selection with
determinants of prior selection (Camron and Heckman, 1998). Unfortunately, like most analysts who do not implement full structural models, we are unable to control for selection prior to high school graduation.

In two empirical papers, Heckman and his associates (1997; 1998) argue that the emphasis in recent econometric literature on the elimination of selective differences in unobservables is misplaced. Showing that selection bias (on unobservables) is a relatively small part of bias as conventionally measured, they conclude that simple balancing of observables in the treatment and control group samples goes a long way toward producing a more effective evaluation strategy. However, it is unclear whether this conclusion can be generalized to other situations.

The most used method, but also the most criticized, is Instrumental Variable (IV) correction. IV techniques have three main weaknesses (Winship and Morgan, 1999). First, assumptions that exclusion restrictions are valid are generally untestable. Second the standard errors of IV estimates can be large if the instrument is weak. Third, IVs only consistently estimate the true average treatment effect when the treatment effect is constant for all individuals. In our case, we believe all three problems exist. Moreover, Heckman et. al (1997) show that using IV to assess the benefits of program participation produces substantially biased estimators of program impact especially when persons do have private (thus unobserved by the analyst) information that is useful for forecasting the gains from participating in the program. In our case, using IV would require us to
Assessing the Mismatch Hypothesis

assume ignorance or irrationality on the part of students. Evidence presented in Heckman et al., (1998) also does not justify application of IV. Therefore we choose not to use this technique.

Six institutions were excluded from the analysis based on the C&B data: four historically black colleges and universities (HBCU) and two universities that did not provide the detailed information needed to measure the timing of graduation. The HBCUs (all are classified as non-selective schools) are not included in the C&B analysis because their selectivity level is lower than that of the other C&B institutions (classified as very, highly and most selective). HBCUs are included in the analyses utilizing the NELS:88 and HS&B datasets in the non-selective category where they belong.

For most institutions, the C&B data files included the entire entering cohorts. However for some institutions the data are derived from samples (Bowen and Bok, 1998). In these cases, sample weights are equal to the inverse of the probability of being sampled. All descriptive statistics presented use appropriate sample weights so that the results accurately represent the entire entering cohort at each institution. These weights allow projections to all C&B institutions (but not to the entire postsecondary universe).

This correction affects the estimated standard errors and the variance-covariance matrix of the estimators, but not the estimated coefficients.
Assessing the Mismatch Hypothesis

10 Using this strategy, a modified zero-order method, we fill all missing data with zeros and add a dummy variable that takes the value one for missing observations and zero for complete ones. These flags provide a useful method for testing whether the pattern of missing observations is random with respect to Y. The modified zero-order strategy is the simplest solution when the proportion of missing data is small (Anderson, Basilevsky and Hum, 1983). We find that the missingness pattern is random in all three datasets, namely it is not related to either the graduation likelihood or to the selection into top institutions. However, as with most remedies for missing data, it does not completely eliminate its potential biasing effects.

11 That Asians’ graduation probability is lower in 1992 than in 1982 may reflect the greater heterogeneity of the Asian sample in 1992 compared with 1982. The earlier sample includes a higher share of South and East Asian youth, who are known to have very high rates of academic success. In addition, their high graduation rate in 1982 may be inflated due to the small sample size, which increases the error of the estimate.

12 For the C&B we report the percent in the top decile of their class because of the restricted variation on this item set for highly selective institutions.

13 We use the STATA –nnmatch- command that implements nearest-neighbor matching estimator for average treatment effect (Abadie et al., 2004). This command corrects for bias associated with matching on multidimensional covariates, as in our case.
For a dummy variable the marginal effect represents the change in the probability associated with a discrete change in the variable from 0 to 1, holding other variables at their mean (Long, 1997).

The null hypothesis is \( \rho(\varepsilon, \nu) = 0 \). That the \( \rho \) coefficient is not significant suggests we cannot reject this null hypothesis that \( \varepsilon \) and \( \nu \) are not correlated. In other words, we cannot reject the hypothesis that there is no selection on unobservables. However, the corresponding standard errors of the \( \rho \) coefficient are quite large along with wide confidence intervals. Such wide confidence intervals containing the null hypothesis value of the parameter signify the lack of precision in our inference. Overall, this suggests that the data is not informative enough regarding the existence of selection on unobservables in the HS&B and NELS:88.
Table 1
Racial and Ethnic Composition of Post-secondary Institutions by Admission Selectivity Tier and Entry Cohort

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-selective</td>
<td>Selective</td>
<td>Non-selective</td>
</tr>
<tr>
<td>White</td>
<td>82.5</td>
<td>85.4</td>
<td>79.4</td>
</tr>
<tr>
<td>Black</td>
<td>11.4</td>
<td>5.3</td>
<td>10.3</td>
</tr>
<tr>
<td>Hispanic</td>
<td>4.4</td>
<td>4.4</td>
<td>6.2</td>
</tr>
<tr>
<td>Asian</td>
<td>1.7</td>
<td>4.9</td>
<td>4.1</td>
</tr>
<tr>
<td>%</td>
<td>83.2</td>
<td>16.8</td>
<td>72.2</td>
</tr>
<tr>
<td>Mean SAT</td>
<td>918 (171)</td>
<td>1082 (174)</td>
<td>911 (167)</td>
</tr>
<tr>
<td>N</td>
<td>3,847</td>
<td>857</td>
<td>3,172</td>
</tr>
</tbody>
</table>
Table 2
Graduation Rates by Race, Institutional Selectivity Tier and Entry Cohort

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-selective (1)</td>
<td>Selective (2)</td>
<td>Odds Ratio&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>White</td>
<td>53.4</td>
<td>82.0</td>
<td>3.9</td>
</tr>
<tr>
<td>Black</td>
<td>26.4</td>
<td>51.6</td>
<td>2.7</td>
</tr>
<tr>
<td>Hispanic</td>
<td>25.7</td>
<td>62.1</td>
<td>2.9</td>
</tr>
<tr>
<td>Asian</td>
<td>50.9</td>
<td>90.9</td>
<td>10.4</td>
</tr>
<tr>
<td>Average</td>
<td>49.0</td>
<td>80.0</td>
<td>3.8</td>
</tr>
</tbody>
</table>

N = 3,847, 857, 3,172, 1,358, 22,253, 6,765

<sup>a</sup> All the odds ratio are statistically significant at P ≤ .01

OR= (odds of success/odds of failure at selective inst.)/(odds of success/odds of failure at nonselective inst.)
### Table 3
The Mismatch Regime: Deviations from Mean Institutional Scholastic Achievement and Class Rank by Selectivity Tier, Entry Cohort and Race

<table>
<thead>
<tr>
<th>Tier: Non-selective</th>
<th>Selective</th>
<th>Tier SAT mean</th>
<th>(s.d.)</th>
<th>Tier class rank mean</th>
<th>(s.d.)</th>
<th>Deviations from Tier Means:</th>
</tr>
</thead>
<tbody>
<tr>
<td>HS &amp; B, 1982</td>
<td></td>
<td>930.7</td>
<td>1112.6</td>
<td>64.6</td>
<td>78.5</td>
<td>White</td>
</tr>
<tr>
<td>NELS, 1992</td>
<td>Selective</td>
<td>912.7</td>
<td>1090.2</td>
<td>67.4</td>
<td>81.4</td>
<td>Black</td>
</tr>
<tr>
<td>C &amp; B, 1989</td>
<td>Selective</td>
<td>1185.3</td>
<td>1333.2</td>
<td>63.7 a</td>
<td>83.8 a</td>
<td>Hispanic</td>
</tr>
<tr>
<td></td>
<td>Most</td>
<td>149.2</td>
<td>121.8</td>
<td>2.3 b</td>
<td>1.0 b</td>
<td>Asian</td>
</tr>
<tr>
<td></td>
<td>selective</td>
<td>(182.2)</td>
<td>(179.3)</td>
<td>(25.5)</td>
<td>(19.8)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(176.7)</td>
<td>(172.0)</td>
<td>(23.1)</td>
<td>(18.3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(149.2)</td>
<td>(121.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **White SAT mean**
  - HS & B, 1982: 21.8
  - NELS, 1992: 21.8
  - C & B, 1989: 13.2

- **White Class rank mean**
  - HS & B, 1982: 1.5
  - NELS, 1992: 1.0
  - C & B, 1989: 2.3 b

- **White N**
  - HS & B, 1982: 2680
  - NELS, 1992: 2387
  - C & B, 1989: 17983

- **Black SAT mean Deviation**
  - HS & B, 1982: -180.0
  - NELS, 1992: -155.1
  - C & B, 1989: -170.7

- **Black Class rank mean Deviation**
  - HS & B, 1982: -9.5
  - NELS, 1992: -6.9
  - C & B, 1989: -25.3 b

- **Black N**
  - HS & B, 1982: 549
  - NELS, 1992: 298
  - C & B, 1989: 1770

- **Hispanic SAT mean Deviation**
  - HS & B, 1982: -115.2
  - NELS, 1992: -100.1
  - C & B, 1989: -83.3

- **Hispanic Class rank mean Deviation**
  - HS & B, 1982: -10.6
  - NELS, 1992: -0.1
  - C & B, 1989: -11.9 b

- **Hispanic N**
  - HS & B, 1982: 467
  - NELS, 1992: 278
  - C & B, 1989: 861

- **Asian SAT mean Deviation**
  - HS & B, 1982: 7.4
  - NELS, 1992: -19.3
  - C & B, 1989: 49.7

- **Asian Class rank mean Deviation**
  - HS & B, 1982: 9.1
  - NELS, 1992: -5.7
  - C & B, 1989: 2.8 b

- **Asian N**
  - HS & B, 1982: 151
  - NELS, 1992: 209
  - C & B, 1989: 1639

---

a) For C&B data, classrank represents percent ranked in top decile.
b) The number reported is the deviation from the corresponding tier average percentage in HS top ten percent.
### Table 4
Group-Specific Marginal Effects (Discrete Change)\(^a\) of Attending a Selective Institution on 6-Year Graduation Status

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline graduation probability(^b)</td>
<td>PS(^c)</td>
<td>ME(^d)</td>
</tr>
<tr>
<td>All Students</td>
<td>0.51</td>
<td>0.18 **</td>
<td>0.18 **</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Group-Specific Models</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0.56</td>
<td>0.17 **</td>
<td>0.19 **</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>0.36</td>
<td>0.14 *</td>
<td>0.18 **</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.38</td>
<td>0.17 **</td>
<td>0.12 *</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>0.50</td>
<td>0.31 **</td>
<td>0.36 **</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) P ≤ .05; \(^b\) P ≤ .01; † P ≤ .10

a) The marginal effect represents the change in the probability associated with a discrete change in the variable from 0 to 1, holding other variables at their mean (Long, 1

b) Predicted probability of graduating from a nonselective institution controlling for family background and academic achievements.

c) Propensity Score results.

d) Matching Estimator results.

e) Marginal effects are reported only when the rho coefficient was found to be statistically significant.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>HS &amp; B</th>
<th>(Std. Dev.)</th>
<th>NELS</th>
<th>(Std. Dev.)</th>
<th>C &amp; B</th>
<th>(Std. Dev.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grad6</td>
<td>6-year graduation proportion</td>
<td>0.54</td>
<td></td>
<td>0.65</td>
<td></td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>White, not of Hispanic origin</td>
<td>0.83</td>
<td></td>
<td>0.79</td>
<td></td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>Black, not of Hispanic origin</td>
<td>0.10</td>
<td></td>
<td>0.09</td>
<td></td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>Hisp</td>
<td>Hispanic, regardless of race</td>
<td>0.04</td>
<td></td>
<td>0.06</td>
<td></td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Mexican</td>
<td>Hispanic of Mexican origin</td>
<td>0.45</td>
<td>0.47</td>
<td></td>
<td></td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Puerto Rican</td>
<td>Hispanic of Puerto Rican origin</td>
<td>0.13</td>
<td>0.11</td>
<td></td>
<td></td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Cuban</td>
<td>Hispanic of Cuban origin</td>
<td>0.42</td>
<td>0.42</td>
<td></td>
<td></td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>Asian or Pacific Islander</td>
<td>0.02</td>
<td>0.06</td>
<td></td>
<td></td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>Selective</td>
<td>If attend a selective institution (most selective in C&amp;B)</td>
<td>0.17</td>
<td>0.28</td>
<td></td>
<td></td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>Thresholds for selective institution</td>
<td>HS&amp;B: Selective :above 1050; Most Selective: above 1250</td>
<td>C&amp;B: Selective :above 1050; Most Selective: above 1250</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parents, B.A.</td>
<td>At least one parent with a BA degree</td>
<td>0.19</td>
<td>0.22</td>
<td></td>
<td>0.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parental income</td>
<td>In categories (HS&amp;B:6; NELS:15; C&amp;B:17 )</td>
<td>0.10</td>
<td>0.10</td>
<td></td>
<td>0.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class rank</td>
<td>HS class rank in percentile (C&amp;B: in Top 10%)</td>
<td>50.01</td>
<td>(36.20)</td>
<td>55.19</td>
<td>(35.79)</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>SAT</td>
<td>SAT score</td>
<td>960.47</td>
<td>(193.83)</td>
<td>963.25</td>
<td>(192.81)</td>
<td>1217.03</td>
<td>(156.07)</td>
</tr>
<tr>
<td>Athlete</td>
<td>If student was an athlete</td>
<td>0.12</td>
<td>0.09 (proxy)</td>
<td></td>
<td>0.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>Female=1, Male=0</td>
<td>0.53</td>
<td>0.55</td>
<td></td>
<td>0.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Public</td>
<td>If attend a public HS</td>
<td>0.83</td>
<td>0.75</td>
<td></td>
<td>0.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>If home region in South</td>
<td>0.29</td>
<td>0.31</td>
<td></td>
<td>0.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Midwest</td>
<td>If home region in Midwest</td>
<td>0.30</td>
<td>0.27</td>
<td></td>
<td>0.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>West</td>
<td>If home region in West</td>
<td>0.14</td>
<td>0.15</td>
<td></td>
<td>0.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>Residing in a rural area</td>
<td>0.28</td>
<td>0.27</td>
<td></td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td></td>
<td>4,704</td>
<td></td>
<td>4,530</td>
<td></td>
<td>29,018</td>
<td></td>
</tr>
</tbody>
</table>