

The Effect of Law School Racial Preferences on Minority Bar Passage Rates **Project Description**

I. The Problem

As America moved from policies of racial exclusion to policies of inclusion in the 1960s, many in American higher education hoped that declines in discrimination would produce greater integration and a convergence in social and economic outcomes between blacks and whites. Though this occurred in some settings, it seemed clear to most observers that progress would be very slow if racial barriers to entry were simply dropped without any affirmative efforts to improve access and participation. One obvious strategy involved outreach (Toepfer 1966). But a fundamental obstacle for many schools was the underlying, pervasive gap between blacks and whites on standardized tests and other measures of academic achievement. On most such exams, such as the SAT and the LSAT, black scores lagged behind whites (usually by more than a full standard deviation); this meant that at elite colleges and most graduate programs, admissions based largely on hard credentials would produce only tiny numbers of black enrollees. Many schools consequently adopted policies in the late 1960s and early 1970s that essentially used different quantitative admissions standards for different groups (especially blacks) so that the racial/ethnic composition of the students admitted would more closely reflect the composition of the underlying applicant pool.

These large racial preferences were initially viewed as a temporary expedient (*Bakke*, 1978). A key reason for the preferences was to convey to minorities that elite educations really were open in a way they had not been before; the preference policies did indeed powerfully stimulate a steady growth in the minority applicant pool. Most educators also assumed that as other civil rights reforms took hold (especially reform of K-12 education systems), the black-white test score gap would rapidly decline. It was anticipated that within a generation, racial preferences would be unnecessary to assure adequate representation.

It did not prove so. The black-white test score gap declined significantly in the 1960s and 1970s, but progress largely stopped around 1980, with the gap still on the order of a standard deviation (Jencks and Phillips, 1998). Recent research has shown compellingly many of the environmental influences that make up the gap (Fryer & Leavitt, 2004). Among law school applicants, the gap has continued to narrow, but remains more than a full standard deviation (Sander 2004). The size of the gap has, however, encouraged elite and graduate schools to practice some variation of race-norming applications.

The early 1990s saw the renewal of a national debate on whether large preferences should continue to be standard practice in much of higher education. On the legal and policy front, the debate gathered steam with *Hopwood*, generated national attention with the passage of the California Civil Rights Initiative (Prop 209), and climaxed with the *Gratz* and *Grutter* decisions of 2003, which left education administrators with considerable discretion in their use of preferences to promote

diversity on campuses. Within the academic community, however, the debate has evolved more slowly. Some researchers have examined the extent of preferences while others have studied the impact of a diverse student body (Jencks, 1991, Thernstrom 1997). But a more central debate has concerned the empirical question of whether persons who received a racial preference benefited from it, and if so, how much.

The question of whether racial preferences are helpful or hurtful to individual beneficiaries has been discussed by educators and economists since at least 1970. Many educators have simply equated preferences with access, and assumed preferences were beneficial because the implicit alternative, exclusion, was obviously bad. Others have noted that for many (if not most) minorities, preferences did not determine whether they could go to college or professional school at all, but rather where in the “eliteness” hierarchy they would land. Landing higher up in this hierarchy is generally thought to be a good thing, for several reasons. A more elite degree seems to carry more weight in the job market (or in subsequent stages of education). At more elite institutions, students make more powerful connections. More elite schools tended to have more resources per student and more talented professors (at least in some dimensions), and this should help students. And, last but not least, the actual immersion in an environment of more talented students is arguably a stimulus to higher achievement – an advantage, not a disadvantage (Winston 2003).

Opposing this reasoning is the theory of “educational mismatch”, articulated by early observers such as Thomas Sowell (1974) and Clyde Summers (1970), and discussed by many later critics of preference policies (e.g., Thernstrom 1997). The mismatch theory really involves two different types of argument. The first is that optimal learning will not occur if a student’s own preparation and cognitive skills are substantially less developed than those of most of the student’s peers, especially in competitive, fast-paced settings. Teachers may pitch their instruction at a level too difficult for the mismatched student to fully absorb; difficulty keeping up becomes more of a concern as the semester wears on, and a student will ultimately learn less and perform badly. The effect is similar to a college freshmen trying to skip first-year physics and go straight into advanced classes: some students might learn faster, but many would crash and burn. (Note that this effect does not turn on race per se; any student facing a large credentials gap with his classmates is vulnerable.)

The second part of the mismatch argument focuses on the group dynamics that may result from the creation of an environment in which entering credentials correlate strongly with race. If credentials even moderately predict school performance, then performance (e.g., grades) will correlate with race as well. This may have a variety of harmful effects: reinforcement or creation of negative stereotypes among whites, feelings of victimization among minorities, tendencies towards group self-segregation, and so on.

Although these various hypotheses have often been raised and discussed, there was little more than anecdotal evidence behind any of them until the intense public discussion of higher education preferences that arose in the mid-1990s. Loury and Garman (1996) used the High School and Beyond database to explore how relative

credentials affected performance. They found evidence consistent with the mismatch effect for both blacks and whites. Students with credentials well below their classmates' averages had lower grades, usually lower graduation rates, and usually lower earnings than students with similar credentials but smaller or no mismatch.

A much larger and far more discussed study appeared in 1998. In *The Shape of the River*, Bowen and Bok used specially developed data (known as the College and Beyond database) which described three large, longitudinal cohorts of students who had attended colleges ranging from moderately elite to very elite. Bowen and Bok found little evidence of a mismatch effect. Blacks and Hispanics who attended more elite colleges had higher graduation rates and higher earnings than did students with similar credentials who attended less elite schools. Bowen and Bok did find large grade disparities between blacks and whites (the average black GPA in their sample was at the 22nd percentile), but they suggested that this was due to black underperformance as much as it was due to admissions preferences. Indeed, the authors argued that racial preferences at elite colleges in the 1980s and 1990s were relatively modest in many ways, and that race was generally not a determinative factor in admissions.

Critics like Thernstrom (1999) pointed to some of the problems with the Bowen & Bok analysis – problems which continue to bedevil much research on the effects of preferences. The most serious of these was selection bias. A typical Bowen and Bok regression looking at future earnings, for example, would classify each student as falling within one of three ranges of high school grades and one of four ranges of SAT scores. The regression would then, in effect, compare the long-term earnings of pairs of students with similar credentials (e.g., middle third of their high school class and SAT score between 1200 and 1299) but different colleges (e.g., Harvard vs. University of North Carolina), and find that the Harvard graduate earned more than the UNC graduate. But in truth, the two students were probably not quite comparable. The Harvard student would be more likely to be near the top of the measured credential ranges (e.g., have grades near the top of the middle third and SAT scores close to 1299) than would the UNC student. And even if the students had identical numerical credentials, in all likelihood the Harvard student had other “unobserved” credentials that made him stronger than the UNC comparison student (e.g., a better high school, more AP courses, impressive application essays, etc.). Thus, the regression was biased towards finding better outcomes for Harvard students.

Dale and Kreuger (2004) demonstrated how important these selection effects could be in their reanalysis of the Bowen and Bok data. They found that it was possible, using the College and Beyond dataset, to match pairs of students who had been admitted to the same pair of institutions (e.g., again, Harvard and UNC) but had chosen to attend different schools (one to Harvard and one to UNC). This approach probably avoided selection bias, since the comparisons were between two students who were both strong enough to secure admission to the more elite institution. (Of course, this does not create a true experimental condition since student choices are not made randomly; it is possible that students who choose the more elite institution tend to have other attributes that make them stronger, or weaker, than their counterpart choosing the less elite school. But this at

least vastly reduces selection bias.) Using several variations of this method, Dale and Krueger found either no effect or a negative effect on earnings for more elite students – just the opposite of the results found by Bowen and Bok.

Of course, the main premise of the “mismatch” hypothesis is that graduation rates and earnings are just indicators of a more fundamental measurement: how much students actually learn. Bowen and Bok relied on graduation rates as their primary “academic” outcome measure – and a good deal of the other mismatch literature discusses graduation rates. But graduation is a conspicuously imperfect measure. Each college has complete control over its own graduation rate, and many elite schools make great efforts to graduate all of their students, regardless of the student’s actual academic achievements.

A more direct approach to the mismatch effect was taken by Elliott (1996) and Cole and Barber (2003). In two separate studies, these scholars examined whether admissions preferences might affect the course of study and, by extension, the career trajectories of students. Elliott found that students who had credentials significantly lower than their classmates were far more likely than other students to abandon science majors. The intuition behind the finding is compelling: students with credentials well below average will find large, competitive, curve-graded science classes particularly intimidating, and will drift towards “softer” curricula. Cole and Barber, approaching a different question with different data, obtained a very similar result: black students who faced a credentials gap with their classmates were much less likely to stick with an early interest in academia. The intuition is, again, straightforward: students are less likely to want to become professors if they are not fairly successful relative to their college peers.

II. Law School

Until recently, nearly all the empirical analyses concerning the effects of preferences (including those discussed above) had focused on undergraduate education. Sander’s (2004) analysis of law school data broadened the discussion to graduate school, and drew more general attention than any other work on affirmative action since Bowen and Bok (1998), both because he found evidence of substantial mismatch effects and because of the unusual data at his disposal. In the early 1990s, the Law School Admissions Council (LSAC) began a large-scale longitudinal study of the cohort entering law school in 1991. The LSAC’s purpose was not to study the effects of preferences per se, but to determine the extent to which bar exams produced disparate outcomes. This Bar Passage Study (BPS) thus focused on gathering information on pre-law credentials, law school performance and experiences, and bar outcomes, which were tracked for three years after graduation. With some ninety percent of all accredited American law schools participating, and some seventy percent of the students at these schools participating in the study, the BPS dataset was large (about 27,000 observations) and reasonably representative.

The BPS thus provided the means to create a richer documentation of affirmative action in a specific realm of higher education than had previously been possible . Setting

aside a handful of historically black law schools, Sander showed that that gap in entering credentials between blacks and whites was very large at all law schools, regardless of a school's eliteness. The black-white gap in credentials at a typical law school generally appeared to be about two standard deviations or more (Sander 2004). This pattern was most easily explained by schools using different admissions standards for different racial groups (which is sometimes called "race-norming"). Moreover, incoming credentials proved to be unbiased predictors of performance in law school (i.e., black underperformance, contrary to Bowen and Bok's findings for college, was either nonexistent or minimal in law school¹); the large racial gap in credentials at each school thus translated into a large gap in grades. Again excepting historically black law schools, the median black at nearly all other schools obtained grades at the 6th to 8th percentile of the white grade distribution. In other words, half of all black students ended up in the bottom tenth of their class.

A central question raised by these patterns is whether the low grades blacks undeniably incurred as a result of preferences are offset by the value of attending a more elite school. Which is more important, an elite school or good grades, in graduating from law school, passing the bar, and getting a good job? One could try to answer this question by seeing whether, when one controls for entering credentials, black law students are more or less likely to have good outcomes when they graduate from more elite schools – essentially the Bowen and Bok approach. The problem with this approach is the familiar one of selection bias; if one compares two blacks with LSATs of 155 and UGPAs of 3.4, it is likely that the student going to the more elite school is there for a reason – there are likely to be unobserved characteristics (e.g., attendance at a more elite college, following a tougher curriculum, demonstrating stronger writing ability) that make the more "elite" student academically stronger, independent of law school effects. The problem is greatly compounded by the fuzziness with which the BPS database measured eliteness. Rather than identifying individual schools, BPS classifies them into one of six "clusters", which are organized partly, but not exclusively, by school eliteness. It would appear, for example, that Vanderbilt Law School is classified in what is, on the whole, the most elite of the six clusters, while the University of Michigan Law School is classified in the second-most elite cluster, though by nearly any school-by-school ranking, Michigan would rank above Vanderbilt). For all these reasons, a simple regression on individual credentials would produce fuzzy and non-significant results.

To deal with these weaknesses in the data, Sander used a more deductive argument: if students receiving preferences essentially trade good grades for a more elite degree, which is more important in determining desirable outcomes? For both graduation and bar passage outcomes, the BPS measures of grades were far more powerful predictors than the measures of eliteness (for earnings, grades appeared to be more powerful at all but the most elite schools). Sander estimated that the eliteness/grades tradeoff was bad enough to account for roughly half of the black-white gap in bar passage rates; to put it differently, maximizing one's racial preference could roughly double a

¹ Research in the 1970s had found significant evidence of black underperformance in law school, but a comprehensive analysis by Anthony and Liu (2004) found "very slight" black underperformance for more recent cohorts.

black law student's chances of eventually failing the bar. Of course, this line of reasoning was somewhat undermined by the fuzzy measure of school eliteness in the BPS; although grades seemed to be far more powerful than eliteness in predicting outcomes, this might be due to the far greater precision with which the database measured grades.

In a subsequent analysis, Sander (2005) used a different methodology to examine outcomes with the BPS data. A student questionnaire in the BPS asked matriculating law students about the process of applying to law school. About one-seventh of the black students who had been admitted to more than one law school reported they had turned down their "first choice" school to attend an apparently less elite school, usually for financial or geographic reasons. Examining black law students who had been admitted to a more elite school but had chosen a less elite one (known as "second choice" students) provided a way – similar in some respects to the Dale-Krueger method – of avoiding the problem of selection bias. With this method, there was a greatly reduced risk that differing outcomes between the two groups could be attributed to unobserved differences in incoming credentials. The technique also mooted complications of race, because one could compare blacks directly with other blacks. The results of this analysis were highly consistent with Sander's earlier estimates. Blacks choosing to attend less elite schools had significantly higher grades and reduced their chance of failing the bar by nearly half - strong results that held up with a wide variety of added controls.

Overall, three alternative views emerge from the literature and discussion of the law school mismatch research: (1) evidence of mismatch is substantial and persuasive (Sander, 2004 and 2005); (2) data is consistent with mismatch, but there may be other, undefined explanations (Rothstein and Yoon 2006 (mismatch may exist for 80% of blacks at non-elite schools, but data problems preclude definitive answer)); (3) data is inconsistent with mismatch, but there is no good alternative explanation of the black-white disparities in outcomes (Chambers et al. 2005 (emphasizing inconclusiveness of intra-black population regressions)).

III. The Need for Better Data

Despite many disagreements over the extent of the mismatch effect, all parties agree that better data is essential if real progress is to be made in understanding the reasons why blacks and Hispanics have so much difficulty with the bar (Rothstein & Yoon 2006; Barnes 2006; Sander 2005) The major shortcomings of the BPS are these:

1) The BPS data provided information only on whether students passed the bar exam on each attempt; it provided no information about actual bar scores. Knowing only dichotomous results means that one is measuring outcomes very crudely (especially if, as in the BPS, seven-eighths of all bar-takers pass on their first attempt). Actual bar scores allow for standard OLS techniques and dramatically better measurement of effects.

This is particularly true in examining mismatch effects at the “top of the distribution.” Some scholars have argued that mismatch effects are non-existent or “reversed” for beneficiaries of preferences at the most elite schools (Chambers, 2005; Ayres & Brooks, 2005). But such a conclusion could follow fallaciously from dichotomous data: if the racial gap in bar scores is identical at elite schools but the absolute scores are higher, the difference in bar passage rates would at some point decline into statistical insignificance.

2) The BPS data on bar results does not identify individual states, but rather combines states into regions. There is often considerable variation in bar passage rates within regions because of the extremely large differences among states in the score they require for passing. In his original analysis, Sander found that this grouping did not bias overall measurement of the black-white gap in bar passage rates (Sander 2004). But there is no question that the lack of individual-state data substantially adds to the indeterminacy of analyses with BPS data, since “passing the bar” means something different in each state and it is impossible to confine a BPS analysis to a single state.

3) As discussed earlier, the BPS data does not allow a researcher to identify individual schools, but only one of six sizeable and often heterogeneous “clusters”. This greatly and intrinsically limits one’s ability to control for a variety of school effects.

4) The BPS covered the cohort of students who completed law school in 1994, but conditions have changed significantly since that time. The summer of 1994 turned out to be the high-water mark for national bar passage rates; many states have significantly toughened pass standards since then (Merritt 2001; Jones 2006). Data on how this has affected blacks is fragmentary, because California is the only state that regularly releases bar results broken down by race. However, triangulation based on recent data from three states (Colorado, Texas and California), combined with national data on graduation rates, strongly suggests that only 33% of blacks who entered law school in 2001 graduated and passed the bar on their first attempt (down from 45% in the BPS data), and only 47% eventually became lawyers (down from 57% in the BPS data). This means that patterns from the 1990s may not apply in the 2000s, and of course it means that the problem is more serious than ever.

The key to creating a richer dataset has been to secure the cooperation of state bar commissions, which administer exams and maintain custody of records.

IV. Our Proposal

Our proposal has two different aspects and goals. The first is to use California bar data to conduct four different types of tests of the mismatch effect – all of them of unprecedented rigor. The second is to complete the first stage of building a multi-state data set that will facilitate more comprehensive and generalizable analyses of issues related to law school preferences and possible mismatch effects within the next two to three years. We will discuss each of these in turn.

A. Analyses of California Bar Data

For understandable reasons, state bar associations have traditionally been highly protective of data on bar scores. The bar organizations are fairly autonomous, they have a strong fiduciary obligation to bar-takers, and most have not seen research as a key part of their mission. Observers thought BPS scored a major coup in the 1990's in getting many states to provide even pass-fail data anonymously and grouped by region.

This picture is starting to change in important respects. With the raising of bar passing scores in several states during the past decade, and concomitant decline in bar passage rates, state bars have been under greater pressure to demonstrate the validity of their exams. Mismatch research has also gotten the attention of bar authorities. For decades, the enormous racial disparity in bar outcomes has been a painful and seemingly insoluble problem for state bars. The mismatch theory implies that law schools may be able to adopt a variety of strategies that could greatly reduce racial disparities. This is a powerful incentive for state bars to undertake close research.

California is the ideal place to begin this work. California is, of course, the most populous state by far, and the State Bar of California (SBC) is second only to New York in membership and the number of test-takers each year. California has the most racially diverse bar in the country; nonwhites make up nearly 40% of those sitting for the bar. The SBC has been the only bar in the country to regularly publish detailed results by race, and it has been the most progressive bar in its pursuit of in-depth research, including studies of the bar's internal validity and of ways in which the bar can better test skills involved in legal practice (one series of such studies led to the introduction of the so-called "performance test" section of the bar exam in the early 1980s, which has since been adopted in some form by almost all other states).

As we will detail below, much of this research has been done by Dr. Stephen Klein. Klein has conducted dozens of SBC studies over the past thirty years on issues ranging from internal validity to ultimate passage rates. He is not only intimately familiar with the data collection processes of SBC, but actually has most of its data in his possession and can vouch for its quality. We are in an advanced state of negotiations with the SBC to conduct the analyses described below, some of which involve the collection of additional, supplemental data. (We are confident of a final decision from the SBC before NSF completes its evaluation of this proposal.)

We expect to obtain authorization to use three SBC datasets. The first, which we will call SBC-1, consists of existing bar data for the exam cohorts from 1997 through 2005. These data include the raw and standardized scores of bar takers on the various components of the bar, the LSAT scores of most bar takers, their race and gender, and the law school they attended. The second dataset, which we will call SBC-2, consists of existing bar data collected by SBC for past validity studies. This dataset covers fewer years, but includes (in addition to the data in SBC-1) data on the law school grades and undergraduate grades of bar takers. The third dataset, which we will call SBC-3, would duplicate SBC-2 for the two most recent cohorts of California bar takers (those taking the

bar in 2004 and 2005), and would include for nearly all bar takers their detailed bar scores, law school attended, race and gender, law school GPAs, undergraduate GPAs, undergraduate college LSAT scores. Only SBC-3 requires significant new data collection efforts. With this data, we will conduct the four analyses described below:

Analysis 1: An analysis of the impact of Proposition 209 on minority bar passage. Natural experiments are common in the social sciences when it is not feasible to conduct true experiments, such as when the focus is on the effects of new laws or policies. The adoption of Prop 209 in California in 1996 provides such a natural experiment. Prop 209 generally prohibited the use of race in state programs, including university admissions. Four California law schools were affected: Boalt, UCLA, UC Davis, and UC Hastings. Boalt and UCLA are both elite schools that had used preferences aggressively prior to the passage of Prop 209. UC Davis also used preferences, but to a lesser degree, and UC Hastings still less. The differences in the proportion of minority students at the elite schools before and after the passage of Prop 209 suggested that this legislation resulted in denying admission to roughly fifty blacks and fifty Hispanics annually at these schools who would have been admitted in the pre-209 regime of large preferences. As a practical matter, those denied admission to these two schools could either attend less elite schools within California or attend Stanford (a private elite school in California) or similarly elite public or private schools out-of-state. To the extent that they chose to remain in-state (and did not go to Stanford), they would have presumably attended a law school where their admissions credentials were more like those of their classmates. It is also plausible that these students would have produced some cascade effect among California schools. For example, if the University of Southern California (USC), which usually loses its best applicants to UCLA and Boalt, in 1997 faces a swelling of more qualified minority applicants, it would presumably use smaller preferences and would reject some of the lower-credentialed minority candidates it would otherwise normally admit. These students would then face a similar choice of leaving the state or attending a lower-ranked school.

The California bar data tracks both in-state and out-of-state applicants; each year, there are several thousand first-time bar takers from California law schools and about 1500 candidates from out-of-state schools. (Over the three years before and after Prop 209, about 1200 in-state blacks and 600 out-of-state blacks took the bar exam, along with a slightly higher number of Hispanics.) Comparing the in-state and out-of-state bar-taking populations before and after the implementation of Prop 209 provides an excellent opportunity to test mismatch effects.

The hypotheses:

(1) Proposition 209 went into effect for law schools in the matriculation year 1997 – students who would normally graduate and take the bar in 2000. We predict that for students taking the California bar from out-of-state law schools, there should be no change in bar scores before or after the passage of Prop 209 (there might be a slight improvement in the average quality of out-of-state students taking the exam, but that could be easily controlled for with data on student LSAT scores).

(2) We predict that for in-state blacks and Hispanics, we will observe the following beginning in 2000: (a) an overall decline in the average LSAT of in-state blacks and Hispanics taking the bar; (b) a decline in the LSAT gap between individual blacks and the median LSAT of students at their law school, and between individual Hispanics and the median LSAT of students at their law school; (c) better overall performance in bar scores among in-state blacks and Hispanics relative to similar out-of-state blacks (controlling for LSAT); (d) higher bar scores for black and Hispanic students, controlling for LSAT, when that student has a smaller credential gap with classmates.

This analysis uses data already on hand (SBC-1). The approach we describe nicely avoids the selection bias problems so pervasive in mismatch studies because the decline in credential differences is an average decline observed for an overall population, not an inferred individual decline based on a smaller test score gap. In other words, we have more confidence that smaller observed credential gaps between blacks and Hispanics vs. their classmates in California after Prop 209 are in fact the result of smaller preferences, compared with the pre-209 and the out-of-state populations.

Analysis 2: Cross-Section Analysis of Bar Score Outcomes using “Case Control” Method. A second basic approach to examining the impact of preferences is through the direct comparison of students with differing levels of credential disparities with their classmates. The straightforward hypothesis is that, other things being equal, we will observe worse bar scores for students (regardless of race) who have a large credentials gap with their classmates than when the gap is smaller.

In Analysis 2, we will construct a logit model to predict a student’s likelihood of passing the bar exam based on that student’s admission credentials (i.e., LSAT score and UGPA, with the later variable adjusted for undergraduate school attended). We will then form pairs of students where both members of the pair had comparable predicted probabilities of passing, but one graduated from a law school where his/her classmates had similar admissions credentials while the other member of the pair graduated from a law school where his/her classmates had a substantially different “credential gap” (i.e., much better or worse). We will then examine whether the difference in credential gap is related to differences in bar scores between members of a pair. This method of analysis corresponds to the type of “case control” study Klein, Freedman, and their colleagues used in their recent study of possible racial bias in the federal death penalty (2006).

There is some room for selection bias in this approach; in comparing a pair of students with similar entering credentials, is the one who is at the more elite institution (and thus potentially more mismatched) really comparable to the other student, or is he at the more elite institution because of some unobserved advantage? However, the bias cuts against a finding of mismatch effects. More to the point, we use three credential factors to compare student credentials: LSAT, undergraduate GPA, and undergraduate college. This third factor – college quality – has probably been the most important unobserved biasing factor in earlier analyses, so we believe selection bias will be quite small. And, of course, this test will be far more powerful than earlier tests in three other respects: the direct (rather than inferred) measure of credential gaps for each student, the use of actual

bar scores (rather than pass/fail outcomes), and the elimination of confounding effects from using multiple bars.

Analysis 3: Cross-Sectional Analysis of Bar Score Outcomes with Regression Techniques. This analysis will be similar to Analysis 2 in all respects except for methodology. Analysis 2 uses a case-control method to pair comparable students. Analysis 3 uses standard regression techniques to compare all students in the population for whom we have complete data. The dependent variable in this model is a bar-takers' exam score. The independent variables include the student's credentials, gender, race, year, and the measured credentials gap between that student and his classmates (i.e., a measure of mismatch).

Analyses 2 and 3 will use two databases, SBC-2 and SBC-3. As noted earlier, SBC-2 includes detailed data on student credentials gathered by the Bar for earlier validity studies. SBC-3 will gather comparable data for the 2004 and 2005 bar-taking cohorts. In this process, we will enlist the cooperation of California law schools (there are nineteen ABA-accredited law schools in the state, and a number of unaccredited schools), using the confidentiality methods described in Part VIII, to gather information on student credentials and law school performance.

Analysis 4: Descriptive Analyses of Preference Patterns. The prior three analyses focus on measuring possible mismatch effects among law graduates. This analysis examines instead a range of other issues relevant to a discussion of admissions preferences in legal education. These include: (a) the current extent of racial admissions preferences by California law schools; (b) the predictiveness of law school admissions criteria in eventual bar performance; (c) the relation between admissions criteria and performance in law school, including the question of whether blacks and Hispanics over- or under-perform standard admissions criteria, or whether racial factors are neutral with respect to law school and bar performance.

B. The Development of a National Database on Legal Education Outcomes

The completion of the analyses we describe in the last section would be a major contribution to the literature on the mismatch effect. But results that would be more generalizable to other states require two additional elements: first, data from a broader cross-section of states around the nation; and second, data on application and admissions patterns among law students. The rationale for both elements is straightforward. Data from a wider range of states is necessary to establish that whatever patterns are found in California can be generalized more broadly. California's bar exam has one of the lowest pass rates in the nation,² and the introduction of Prop 209 arguably has effects that make

² Note, however, that California's bar pass rate is unusually low in large part because it allows students from unaccredited schools to take the bar; the state pass rate for graduates of ABA-accredited schools is only about 4% lower than the national average. Note, too, that our use of bar scores as opposed to pass rates allows us to eliminate whether a student "passes" or "fails" from most of our analyses.

the effects of racial preferences harder to study.³ Data on application and admissions patterns makes it possible to better control for selection-bias problems by finding pairs of students who were admitted to the same pair of law schools but chose to attend schools of differing eliteness (the Dale-Krueger method described in Part I, above).

We believe the key to building a better database is to create a foundation on which other institutions can build. We have had discussions with several of the major bars around the country. Like the State Bar of California, they are aware of the mismatch issue, interested in getting to the bottom of the sources of racial disparities in bar outcomes, and desirous of better understanding dynamics within their own states. Their concerns have to do with data confidentiality, cost, and the possible adverse publicity involved in being first down the path of analysis and disclosure. By building a database with the California Bar, by demonstrating the effectiveness of our confidentiality procedures, and producing careful and credible findings, we are confident that other large-state bar organizations will join in our effort. It seems to us very realistic to project that half-a-dozen major bars will be participating in the national database project by the end of this grant in June 2008. With the involvement of state bars will come the participation of many of the law schools within those states, which have traditionally cooperated when state bars have undertaken internal validity analyses.

Another key institutional player in this process is the Law School Admissions Council (LSAC). LSAC administers the entrance examination for law schools (the LSAT) and provides services that facilitate the admissions processes. It generally tracks the schools to which students apply, which ones accept them, and where they initially attend. It also widely collects data on first-year performance in law school for the purpose of conducting validity studies that help schools determine how much weight they should assign to the LSAT (relative to undergraduate grades) in the admissions process. Obviously, this wealth of information would facilitate the research we have discussed. Unlike the state bars – which are generally governed either by the state judicial system or the associated lawyers of the state – LSAC is governed by the accredited law schools. Given the intensely political debate about preferences within law schools, it is not surprising that LSAC has been cautious about undertaking its own research on mismatch issues. Once again, however, we believe that the emergence of a multi-state, contemporary database of bar outcomes will create a very powerful inducement for LSAC to cooperate.

V. Intellectual merit and broader impact

Our study has strong **intellectual merit**. The “mismatch debate” has been one of the most active debates in legal academia in recent memory. Our project brings together several leading researchers in related fields, and creates a database far superior to any ever used to examine the problem at hand. We will analyze this data in sophisticated but very accessible ways, with a range of methodologies. The results of this study will

³ Our proposed analyses do address this issue by including data before as well as after Prop 209, and, of course, in Analysis 1 by explicitly studying the impact of Prop 209.

substantially advance the discourse. And the data we create will greatly facilitate future research.

Our study will also have very **substantial broader impact**. The debate over racial preferences in higher education has been one of the most important and long-lasting in modern America. Legal education – as reflected in the *DeFunis*, *Hopwood* and *Grutter* court decisions – has always near the heart of the debate. Our study will receive national attention and will be an important part of efforts in legal education and higher education generally to make efforts towards racial justice and inclusion both smarter and more effective. These efforts, it is important to emphasize, are not limited to choices about the calibration of admissions preferences. Law schools can also address mismatch problems, to the extent they are found, through curricular reform, academic support, and other methods of better equalizing learning outcomes within institutions.

VI. Project Team

The four principal investigators on this proposal are uniquely well-positioned to carry out this study. Dr. Richard Sander is an economist and law professor at UCLA, and probably the leading author in the country on legal education issues over the past decade. Sander's analysis of law school mismatch effects integrated a wide range of previously unpublished data and has generated an extraordinarily active debate in legal academia.

Dr. Stephen Klein is one of the nation's leading psychometricians and the unquestioned dean of research in bar studies. As a Senior Research Scientist at RAND over the past thirty years, Klein conducted pioneering research on methods of improving the bar's measurement of lawyering skills (which led to the introduction of the "performance" exam in California) and has been a key player in the recent development of the Collegiate Learning Assessment Project, which uses innovative testing methods to assess learning during college, and has been adopted by over two hundred colleges nationwide. Klein's dozens of studies on state bars are considered authoritative, and no one can approach him in the level of confidence the bar organizations have in his work.

William Henderson is an associate professor of law at Indiana University. Henderson brings three important things to the project. (1) His 2004 article on the LSAT exam is widely thought to be the most important and considered recent critique of standardized testing in law school admissions. Henderson thus brings a healthy skepticism about the meaning and interpretation of test results. (2) Henderson is a leader among the younger generation of scholars studying the legal profession. (3) Perhaps most importantly, Henderson has played a prominent national role in encouraging balanced and dispassionate discussion of Sander's research on law school mismatch effects. His involvement in the project will bring credibility to the results among strong supporters of affirmative action and will help insure that a wide range of scholars will examine and discuss our methodologies as we proceed.

Dr. E. Douglass Williams is Chair of the Economics Department at Sewanee University (aka The University of the South). Williams is a cautious, careful labor

economist whose interest in legal education goes back to his PhD dissertation which included a chapter on the role of law schools in the lawyer population explosion during the 1970s. He has a special interest in the market for education and will bring a knowledge of the relevant theoretical literature to the project. He has also been engaged in research using a variety of econometric techniques, including limited dependent variable, methods and will bring an experienced eye for evaluating the econometric methods and results.

The two other staff in the project are Dr. Roger Bolus and Dr. Robert Sockloskie. Bolus received his Ph.D. from U.C.L.A. in Research Methods, Measurement and Evaluation. He is founder and senior partner of Research Solutions Group, a consulting firm that specializes in providing psychometric, statistical analysis, and database management services to large scale research programs. For the past several years, he has played an integral role in supporting several State Bar offices (including the State Bar of California) in the statistical processing of examination results. As a result, he has developed a deep understanding of how law school, LSAC, and state bar data are organized, accessed, and protected. Sockloskie has a Ph.D. in psychology from UCLA, has worked at RAND, and has worked for the past eight years as a data analyst for the Empirical Research Group at UCLA's School of Law. Dr. Sockloskie is a highly-skilled methodologist and statistical analyst.

VII. Past NSF Support

Dr. Sander was co-PI on NSF grant 0115521. *After the JD, Phase I* was funded in 2001; the NSF grant of \$166,000 covered approximately ten percent of the total cost of this longitudinal study of nearly five thousand lawyers who completed law school in 1999 and 2000. After the JD created a weighted national sample of new lawyers, secured background information on the lawyers through LSAC (a co-funder of the project), and secured survey responses or interviews with roughly two-thirds of the designated sample in the third or fourth year of their careers. An initial report on the project was published in September 2004, entitled *After the JD: First Results of a National Study of Legal Careers*. Sander has used data from the study in three recent articles (Sander 2004, 2005, 2006). Part II of the study, which will re-survey participants in the seventh-to-eighth year of their careers, received support in 2006 from NSF's Law and Social Science Program (grant 0550605). Sander is not a co-PI on Part II.

Dr. Klein was a PI on NSF grant ESI-99-86612, which ran from 1999 to 2005 and was entitled *Mosaic II: Longitudinal Investigation of the Effects of Systemic Reforms on Student Achievement*. Mosaic II was a six-year longitudinal research project sponsored by the NSF and co-directed by Dr. Klein and Dr. Laura Hamilton. It was a continuation of RAND's (and Dr. Klein's and Hamilton's) research on the relationship between teacher instructional practices and student achievement in mathematics and science. Results from our earlier NSF studies (grant ESI-96-15809) suggested that exposure to reform-based instruction was only weakly related to student achievement in math and science. Mosaic II allowed for a more in-depth examination of this relationship, including more complete measures of classroom practices, the use of NSF-developed

curriculum materials, and exposure to various staff development opportunities. Mosaic II also investigated the benefits of cumulative exposure to reform practices over three years and the feasibility of using computer-based scoring algorithms to evaluate student responses to open-ended questions. The results of this research culminated in several reports and journal articles, including Vi-Nhuan Le et al (2006), Lockwood et al (in press), Stecher et al (2006), and Hardison et al (under review).

VIII. Confidentiality procedures and Human Subjects Issues

This project does not involve the collection of any new data from individuals, but simply the assemblage of existing data. We described in Part IV (pages 7-8) the three datasets we will use. SBC-1 and SBC-2 are based on existing datasets created by the State Bar of California. Individual identifiers have already been removed from these data. SBC-3 involves the merging of data from the State Bar with data in the archives of California (and possibly some out-of-state) law schools. We propose to do this in the following manner: for each person taking the bar over the period we wish to cover, staff at the State Bar will create two random digit identifiers (call them Random1 and Random2). They will send to Dr. Klein the two random identifier lists, so that each Random1 identifier can be matched against the corresponding Random2. The State Bar will organize the lists by school, and send to each participating school names of its students with the corresponding Random1 identifier. The school will compile data (e.g., law school grades, etc.) for each person in a spreadsheet that contains the student name and Random1 identifier, and then, when data entry and checks are complete, will delete the name column from the spreadsheet. This data will be forwarded directly to Dr. Klein, who will match Random1 identifiers with Random2 identifiers, merge the school data into other data provided by the state bar, and delete the Random1 identifiers from the database. In this manner, only school administrators ever see personally-identified student information (which they, of course, already have access to and work with); only state bar officials sees bar data; the project team will observe no personally-identified information.

The other confidentiality danger posed by a study such as this is that someone using the dataset could deduce an individual's identity through the items included in the dataset (i.e., an Hispanic female who attended Columbia for college and Boalt for law school). To address this concern, we will revise the database to include numerical descriptors rather than proper names of institutions – that is, we will use various statistical information to characterize the eliteness and competitiveness of a particular college or law school and thereby eliminate the use of actual institution names in either our database or our results. We will also suppress data on cohorts at particular institutions with fewer than five members. Only Dr. Klein will retain a complete dataset, for purposes of integrating additional data down the road.

We are in the process of securing appropriate approvals from the UCLA Human Subjects Review Committee.