IS THERE A MISMATCH EFFECT IN LAW SCHOOL, WHY MIGHT IT ARISE, AND WHAT WOULD IT MEAN?

GREGORY CAMILLI*
KEVIN G. WELNER**

INTRODUCTION

A "mismatch" in law school has been defined as the gap between the strength of a student's entering credentials at a particular school and those of the modal student at that school. Accordingly, the mismatch hypothesis concerns a student whose level of entering credentials in a law school falls substantially below the school's average. In particular, the hypothesis stipulates that more learning occurs when a student attends a school where any credentials gap is small, or correspondingly, that less learning occurs when the gap is large.

The credentials upon which the mismatch hypothesis has been framed include undergraduate grade point average as well as the applicant's score on the Law School Admissions Test (LSAT). These presumptively fair and objective criteria are typically implemented in research studies as a weighted combination, and it is often assumed that merit and success in law

school and the legal field are strongly associated with entering credentials.[1] Whether or not this is true, an initial credentials gap is generally created when individuals are admitted to law schools on the basis of preferences outside of those criteria (grades plus scores)—whether due to race, legacy,[2] or some other basis. Proponents of the mismatch hypothesis argue that such students subsequently learn less and consequently have a lower likelihood of becoming practicing attorneys than if they had attended a better-matching school. This hypothesis has gained vocal adherents and detractors since it was first investigated in law school admissions in 2004.[3] The ensuing debate has primarily centered on the role of the mismatch hypothesis in challenging the effectiveness of race-based affirmative action policies. As we discuss below, this is the case even though students' race plays no direct role in this hypothesis; rather, the gap can and does exist for students of all races, and any learning effects likewise exist independently of race.

In this article, we re-examine the existence and nature of the mismatch effect in law school, asking several key questions about this potential effect: its size and direction, its source, and its significance. We begin with a brief history of affirmative action and its legality, followed by a discussion of the efficiency and meritocracy arguments used by opponents of affirmative action. We then place the mismatch hypothesis within that context, considering how and whether it fits as part of those attacks. Next, we examine how race is, and is not, a reasonable part of the mismatch hypothesis discussion, and we describe how match effects can be modeled

---

* Professor, University of Colorado at Boulder School of Education.

** Professor, University of Colorado at Boulder School of Education, and Director, National Education Policy Center.

1. We use the term *credentials* throughout, to mean an attempt to attach an objective measurement to a student based on undergraduate grades and LSAT score.

2. In a recent study of legacy admissions at thirty highly selective colleges and universities, Michael Hurwitz found that legacy status increases the odds of admission by a factor of 3.13. *See* Michael Hurwitz, *The impact of legacy status on undergraduate admissions at elite colleges and universities*, 30 ECON. OF EDUC. REV. 480 (2011), *available at* http://www.sciencedirect.com/. *See also* John Brittain and Eric L. Bloom, *Admitting the Truth: The Effect of Affirmative Action, Legacy Preferences, and the Meritocratic Ideal on Students of Color in College Admissions*, *in* AFFIRMATIVE ACTION FOR THE RICH: LEGACY PREFERENCES IN COLLEGE ADMISSIONS (Richard Kahlenberg ed., 2010). These authors chronicle the enormous advantage of legatees in the admissions process. For example, the number of legacy admits at elite schools is usually 3–4 times the entire number of Black students at the school. *Id.* at 127. They note Justice Ginsburg's point, "The rallying cry that in the absence of racial discrimination in admissions there would be a true meritocracy ignores the fact that the entire process is poisoned by numerous exceptions to 'merit.'" Grutter v. Bollinger, 539 U.S. 306, 367–68 (2003) (Ginsburg, J., concurring).

3. The original article was R. H. Sander, *Systemic Analysis of Affirmative Action in American Law Schools*, 57 STAN. L. REV. 367(2004) [hereinafter *Systemic Analysis*]. A second article central to the present research is R. H. Sander, *Reply: A Reply to Critics*, 57 STAN. L. REV. 1963 (2005) [hereinafter *Reply to Critics*].

and analyzed. Before looking at research concerning those effects, however, we examine the causal assumptions among mismatch hypothesis supporters, and we present research about the effects of heterogeneous (diverse in terms of prior measured achievement) learning environments in related contexts. That is, the hypothesis also raises issues regarding postsecondary and post-baccalaureate/graduate school admissions, and those other levels of education help to contextualize the research in law school admissions. We conclude with some reflections on the policy and practice significance of any mismatch effects, again questioning the linking of these questions to affirmative action debates and suggesting instead that the major implications concern law school instruction and academic supports for entering students.

I. THE CONTEXT: A SHORT HISTORY OF AFFIRMATIVE ACTION POLICY AND LAW

The legality of affirmative action policies has been litigated; as discussed below, the *Grutter* decision from 2003 is the law of the land.[4] Yet future litigation is nonetheless likely, spurred in part by Justice O'Connor's comment in *Grutter*:

> The Court takes the Law School at its word that it would like nothing better than to find a race-neutral admissions formula and will terminate its use of racial preferences as soon as practicable. The Court expects that twenty-five years from now, the use of racial preferences will no longer be necessary to further the interest approved today.[5]

Thus, this issue will probably be re-litigated, in 2028 if not before. In part, then, this article places the mismatch hypothesis within a framework of plausible future litigation. The past history of affirmative action policies and jurisprudence is also instructive merely to increase understanding of the current policy arguments.

Many excellent histories of affirmative action have been written.[6] Here, we present only a short overview. Affirmative action began in earnest during the years immediately following the 1968 assassination of Dr. Martin Luther King. By the 1990s, these programs had become firmly established,[7] and—partly as a result—as of 2005 there were approximately 40,000 Black[8] lawyers in the U.S.[9] But this recent progress stands in stark

4. Grutter v. Bollinger, 539 U.S. 306 (2003).

5. *Id.* at 309-10.

6. *See, e.g.*, CHRISTOPHER EDLEY, NOT ALL BLACK AND WHITE: AFFIRMATIVE ACTION, RACE, AND AMERICAN VALUES (1996); IRA KATZNELSON, WHEN AFFIRMATIVE ACTION WAS WHITE (2005).

7. Henry Ramsey, Jr., *Historical Introduction*, *in* Linda F. Wightman, *LSAC National Longitudinal Bar Passage Study*, *infra* note 64, *available at* http://www.sac.org/LSACResources/Research/RR/Wightman-LSAC-98.pdf.

8. A majority, but not all, Black law students are African American, so the

contrast with a long history of exclusion. Prior to the Reconstruction Era, Blacks were not allowed to receive a legal education, and no Black person was admitted to the American Bar Association prior to 1911.[10] There were approximately 1,300 Black lawyers in 1930; this number had increased to just 2,000 by 1960,[11] and few potential Black lawyers were forthcoming. As noted by Richard Sander, "In 1964, there were only about 300 first-year black law students in the United States, and one-third of these were attending the nation's half dozen black law schools."[12] Few students of color were enrolled in historically White law schools prior to the establishment of affirmative action programs.

The Supreme Court has considered affirmative action plans in each of its main contexts: contracting,[13] hiring,[14] and higher education admissions.[15] As a rule, the Court has applied strict scrutiny to any policy that classifies individuals based on their race, meaning that the policy will be found to violate the equal protection clause of the Fourteenth Amendment unless it serves a compelling purpose and is narrowly tailored to achieve that purpose. Early affirmative action plans were defended as serving a remedial (and compelling) purpose—addressing past discrimination—and

---

former term is used herein.

9. David B. Wilkins, *A systematic response to systemic disadvantage: A response to Sander*, 57 STAN. L. REV. 1915 (2005).

10. Three African Americans were admitted to the ABA in 1911 (William Henry Lewis, Butler Roland Wilson, and William R. Morris). J. CLAY SMITH, JR., EMANCIPATION: THE MAKING OF THE BLACK LAWYER, 1844-1944, 541-42(1993). However, according to Smith, "In 1912, word spread across the ABA that it had admitted three black lawyers. Predictably opposition to their membership was strongly voiced by southerners. . . .[a]sserting that the ABA was a social organization." *Id.* at 542. All three men were pressured to resign, but Wilson and Lewis refused. *Id.* At the ABA's annual convention in 1912, a resolution was adopted requiring that the race of Blacks recommended as members be identified. *Id.* at 543. On the basis of this resolution, it was understood that Lewis and Wilson would retain membership, but that future recommendations of African Americans could be vetoed by ABA board members. *Id.* A third African American lawyer, T. Gillis Nutter, was admitted to the ABA in 1929, but the ABA remained White with few exceptions until 1943 when the ABA amended its by-laws to require four, instead of two, negative votes to deny membership. *Id.* at 544-45. Southern voting strength was accordingly diluted. *Id. See also* Robert V. Ward, *From the Slave Quarters to the Courtroom: The Story of the First African American Attorney in the United States*, BLACKPAST.ORG, http://www.blackpast.org/?q=perspectives/william-henry-squire-johnson-slave-quarters-courtroom.

11. J. CLAY SMITH, JR., EMANCIPATION: THE MAKING OF THE BLACK LAWYER, 1844-1944, 565 (1993).

12. *Systemic Analysis*, *supra* note 3, at 375.

13. City of Richmond v. J.A. Cronson Co., 488 U.S. 469 (1989); Adarand Constructors, Inc. v. Peña, 513 U.S. 1012 (1994).

14. United Steel Workers of Am. v. Weber, 444 U.S. 889 (1979); Wygant v. Jackson Bd. of Educ., 478 U.S. 1014 (1986); Johnson v. Santa Clara Cnty. Transp. Agency, 480 U.S. 616 (1987); Metro Broad., Inc. v. FCC, 497 U.S. 547 (1990).

15. Regents of the Univ. of Cal. v. Bakke, 438 U.S. 912 (1978); Gratz v. Bollinger, 539 U.S. 244 (2003); Grutter v. Bollinger, 539 U.S. 306 (2003).

were generally upheld if they were of limited scope.[16]

The Supreme Court has also endorsed "diversity" as a compelling governmental interest in higher education.[17] Current jurisprudence on affirmative action in higher education admission is based on twin cases concerning the University of Michigan. In *Gratz v. Bollinger*, the Court found the University's undergraduate admissions program to be in violation of the Equal Protection Clause because, as Justice O'Connor explained in the *Grutter* case (decided concurrently with *Gratz*), it implemented rigid, "mechanical, predetermined diversity 'bonuses' based on race or ethnicity."[18] This rule consisted of simply adding points for applicants based on race or ethnicity. While other point increases were also included in the system, racial classification is expressly addressed in the Fourteenth Amendment, so only those elements were subject to the challenge and to strict scrutiny.

In contrast, the Court in *Grutter v. Bollinger* upheld the University of Michigan law school's affirmative action policy, which used race merely as a "potential 'plus' factor" and as part of a larger, comprehensive review of applicants' files.[19] The Court reaffirmed (five to four) that diversity is a compelling state interest and can be tailored to have multiple sources, including "racial and ethnic diversity with special reference to the inclusion of students from groups which have been historically discriminated against."[20]

Where the mismatch hypothesis would fit within the Court's legal framework, should the hypothesis find sufficient empirical support, is not clear. Most likely, a plaintiff challenging an affirmative action policy at a law school would argue on the basis of a mismatch effect that the policy is not narrowly tailored. Even given a compelling interest in diversity, using race to place applicants in an environment where they would be less likely to succeed is a poor approach for pursuing that goal. Secondarily, a plaintiff might use the mismatch effect to reargue the basic idea that diversity in higher education is indeed a compelling state interest.[21]

---

16. *See, e.g.*, Fullilove v. Klutznick, 448 U.S. 448 (1980).

17. Regents of the University of California v. Bakke, 438 U.S. 912 (1978); Grutter v. Bollinger, 539 U.S. 306 (2003).

18. Grutter v. Bollinger, 539 U.S. 306 337(2003).

19. Grutter, 539 U.S. at 307 (2003).

20. *Id.* at 316.

21. Note that a potential empirical demonstration of a mismatch effect for students of color would be primarily evidence that mismatch effects exist for students of *any* race or ethnicity. This is because, as discussed later in this article, the causal assumptions underlying the mismatch hypothesis have nothing to do with race. Thus, the approach would be to show the effect and then only indirectly show that mismatch harms Black students through the correlation of race with admission preference. We should note, however, that using such empirical evidence to implement policy changes solely for Black students would be an arbitrary use of correlative evidence. It would not eliminate mismatch (as defined to date) from the population of law school students.

More remotely, plaintiffs might attempt to use "undue harm" language, such as that set forth in Justice O'Connor's opinion in *Grutter*:

> We acknowledge that "there are serious problems of justice connected with the idea of preference itself." Narrow tailoring, therefore, requires that a race-conscious admissions program *not unduly harm members of any racial group*. Even remedial race-based governmental action generally "remains subject to continuing oversight to assure that it will work the least harm possible to other innocent persons competing for the benefit."[22]

While Justice O'Connor's concern about "undue harm" was focused on those who were not admitted to the elite law school, this language might be extended—if the mismatch hypothesis were to prove empirically grounded—to intended beneficiaries of the policy. Those admitted might be argued to suffer undue harm if they are being provided with an education that undermines their future success.[23] This is, in fact, the apparent thinking behind the policy push surrounding the mismatch hypothesis.

II. MISMATCH, EFFICIENCY, AND MERITOCRACY

The most common argument against affirmative action in law school admission is not the mismatch hypothesis, but rather the efficiency

---

22. Grutter, 539 U.S. at 341 (emphasis added; internal citations omitted).

23. To be clear, we see no merit in this legal contention. In part, the lack of merit is due to evidentiary weakness as outlined later in this article (i.e., the hypothesis has little evidence behind it). But any "undue burden" argument will likely go nowhere as a legal matter because the state action is only acceptance of the applicant—not forcible conscription into law school. Moreover, the benefits of attending an elite law school go beyond bar-exam preparation.

It should also be noted that arguments against affirmative action include a progressive critique pointing to the limited capacity of the approach to address broader structural inequalities in U.S. educational opportunity. Another critique is grounded in the contention that the idea of "diversity" is too general to support the requirements of narrow tailoring and that admissions officers engage in insufficient individual evaluation. Justice Rehnquist's dissent in Grutter mocks the idea that the law school was truly trying to obtain a critical mass of students in different minority groups.

At the state level, a challenge to affirmative action has arisen in the form of referenda prohibiting racial preferences in state hiring and higher education admissions. A number of successful state initiatives have prohibited "discrimination or preferential treatment in public employment, public education, and public contracting," effectively ending affirmative action (such as California's Proposition 209 in 1996; Washington's Initiative 200 in 1998; Michigan's Proposal 2 in 2006; and Nebraska's Initiative 424 in 2008). A similar proposal, Amendment 46, failed to pass in Colorado in 2008. *See* Michele S. Moses, Amy N. Farley, Matthew Gaertner, Christina Paguyo, Darrell D. Jackson, & Kenneth R. Howe (2010). *Investigating the Defeat of Amendment 46 in Colorado: An Analysis of the Trends and Principal Factors Influencing Voter Behaviors*, *available at* http://www.colorado.edu/education/faculty/michelemoses/docs/finalmosesamendment46.pdf.

proposition that performance in law school is mainly based on incoming credentials. This argument, influenced by beliefs in meritocracy and in fixed ability, holds that individual "merit" is the only justifiable selection procedure for two reasons.[24] First, group-based admission criteria are based on the unconstitutional premise that group distinctions can countermand individual suitability, especially with regard to the equal protection clause of the Fourteenth Amendment. Second, scarce resources should be allocated to the most talented applicants, who then return the highest level of benefit to society. This latter approach to selection maximizes utility.[25] Anything else, so the argument goes, should be denounced as a quota system.[26] Students who are given preferential treatment in the admission process are typically less academically prepared than other students, and such differential credentials are asserted to predictably translate into eventual achievement differences.

The traditional argument against affirmative action requires nothing further, despite the presumed equivalence of academic credentials and merit—for which we provide a fuller discussion below. In contrast, this is only the departure point for the mismatch hypothesis. For students entering law school with substantially lower levels of credentials than the average student at a school, the hypothesized mismatch effect is properly understood as a negative net effect beyond an outcome predicted by those credentials. That is, the student is hypothesized to learn less than would be predicted on the basis of credentials. Conceptually, there are two distinct effects—predicted achievement based on entering credentials and the presumed mismatch phenomenon. Interestingly, they tend to be conflated, or possibly simultaneously presented even when one argument might undermine the other. For instance, the quotation below about the purported mismatch effect is from Gail Heriot,[27] a University of San Diego law professor who co-chaired the committee for Yes on Proposition 209 (banning affirmative action in California):

> African-American students attending law schools failed or dropped out at much higher rates than white students (19.3% vs. 8.2%). Overwhelmingly, this phenomenon was associated with

24. *See* ARTHUR R. JENSEN, BIAS IN MENTAL TESTING (1980).

25. Gregory Camilli, *Test Fairness*, in EDUCATIONAL MEASUREMENT (Robert L. Brennan ed., 2006).

26. *See* Jensen, *supra* note 24.

27. Heriot was also appointed by Congress in 2007 to a 6-year term to the U.S. Commission on Civil Rights. Her role in Proposition 209 is described in *California's Proposition 209 and the United States Constitution*, 43 LOYOLA L. REV. 613 (1998). She is also a member of the board of the National Association of Scholars, whose mission is described as follows: "We uphold the principle of individual merit and oppose racial, gender, and other group preferences. And we regard the Western intellectual heritage as the indispensable foundation of American higher education." *Who We Are*, NAT'L ASSOC. OF SCHOLARS, http://www.nas.org/who.cfm (last visited May 24, 2011).

> poor performance and not financial hardship. Since many of these students who left law school would likely have performed better at a less competitive law school they were, in a very real sense, victims of race-based admissions.[28]

This counterfactual proposition ("would likely have performed better") is, as discussed later in this article, at the center of arguments put forward by Richard Sander in the article that initially set forth the mismatch hypothesis, *Systemic Analysis of Affirmative Action in American Law Schools*, as well as in his later response article called *A Reply to Critics*.[29]

But Heriot has also made the efficiency argument, which is the more standard argument against affirmative action in law school admissions. Even if it were shown that Black students performed as well in an elite school as they would have in a non-elite school, she would still oppose affirmative action based solely on the credentials gap:

> Students who attend schools where their academic credentials are substantially below their fellow students' tend to perform poorly. The reason is simple: While some students will outperform their entering academic credentials, just as some students will underperform theirs, most students will perform in the range that their academic credentials predict.[30]

Heriot's argument against affirmative action as *argument in the alternative* might be described as Boolean logic,[31] because it is true if either the mismatch is true *or* if poor performance is due to a relatively lower level of qualification. But if it is true that "most students will perform in the range that their academic credentials predict," as Heriot surmises, then it follows that any mismatch effect will only be at the margins—that it will not change learning or outcomes to any substantial degree.[32]

For this reason, the real importance of any mismatch effect ought not to lie with those opposed to affirmative action. Even if there is an effect, it must be small according to Heriot's logic as well as actual empirical analyses, as discussed below. Rather, the importance of any mismatch effect lies with those concerned about improving the success of those admitted with lower credentials. If there is a mismatch effect, it suggests

---

28. *See* Gail Heriot, *How Mismatches Devastate Minority Students*, MINDING THE CAMPUS, http://www.mindingthecampus.com/originals/2008/03/by_gail_heriot_i_have.html (last visited Oct. 1, 2008).

29. *Systemic Analysis* and *Reply to Critics*, *supra* note 3.

30. Heriot, *supra* note 8.

31. With the "or" operator, a statement is true as a whole if either of the two components are true. While argument in the alternative is a recognized legal strategy, it is sometimes disingenuous as a policy argument.

32. Sander seems to share Heriot's Boolean view of affirmative action. *See Systemic Analysis*, *supra* note 3.

that elite law schools are not doing enough to support such admittees. This argument, as we discuss next, employs a different logic than does the efficiency argument.

III. RACE, ETHNICITY, AND MISMATCH

As noted above, the conjectured harm resulting from mismatch is not limited to students admitted through an affirmative action policy; rather, it would apply to any student who is mismatched for any reason. The mismatch hypothesis is not fundamentally about race or ethnicity, nor is it a hypothesis that requires explicit preferential selection.[33] Rather, it is most accurately described as an intuitive scenario about the prerequisites for and contexts of student achievement. This hypothesis is not bound by race despite the pervasive use of racial and ethnic adjectives such as Asian, Black, Hispanic, and White in the existing literature to describe outcomes of mismatched students.

The question of how race and ethnicity should (or should not) be incorporated into an investigation of the mismatch hypothesis is complex. The current effort to understand the effects of preferential admission is being vigorously pursued, if not led, by Project SEAPHE (Scale and Effects of Admission Preferences in Higher Education), a group of scholars, including Richard Sander, whose goal is to "ground the public's understanding of affirmative action in rigorous, data-driven studies."[34] In particular, these researchers are "seeking [California state] bar records wish to use them to test whether individuals who benefit from admissions preferences perform worse on the bar exam than they would have if they had attended a less elite law school."[35] One activity of the project has been to file suit against the State Bar of California to obtain bar examination data for the purpose of analyzing the effects of admission preferences.[36] The State Bar has refused the SEAPHE request to provide data, based on the argument that the data were not collected from law school applicants for use by third parties or for the purpose of studying the effects of affirmative action.[37] An appeal by SEAPHE was, at the time of this writing, before the California Second District Court of Appeal.[38]

---

33. Katherine Y. Barnes, *Is Affirmative Action Responsible for the Achievement Gap Between Black and White Law Students?*, 101 NW. U. L. REV. 1759, 1808 (2007).

34. *About*, PROJECT SEAPHE, http://www.seaphe.org/about.php (last visited May 24, 2011).

35. *Project SEAPHE Press Kit*, PROJECT SEAPHE, http://www.seaphe.org/pdf/presskit-section1.pdf at 1.

36. Sander v. State Bar of California, S165765 LEXIS 11271 (Cal. 2008).

37. Letter from Gayle E. Murphy, Senior Executive, Office of Admissions, Committee of Bar Examiners of the State Bar of California, TITLE, to Richard Sander, Ph.D., and William Henderson, Professors (July 31, 2007), *available at* http://www.seaphe.org/pdf/bar-proposal/letter_from_murphy.pdf.

38. *See* Sharon L. Browne, *Records on bar exam pass rates aren't exempt from public disclosure*, *available at* http://www.pacificlegal.org/page.aspx?pid=1422. The

Better data would undoubtedly permit more relevant and convincing analyses of the effects of mismatch and would especially help to identify any threshold below which potential mismatch effects become more probable.[39] Actual bar examination scores might resolve effects masked by simple pass-fail measurements on bar performance.[40] Obtaining such data presents problems of anonymity and confidentiality, but there are several convincing arguments that those concerns can be addressed, especially since technical solutions exist for anonymizing sensitive data.[41] However,

---

Court of Appeal is considering arguments based in part on the contention that the Superior Court decision was based on an overly narrow interpretation of California's Proposition 59.

Sander and Joe Hicks, Vice President of Community Advocates, Inc., and the California First Amendment Coalition, originally petitioned the Supreme Court of the State of California to force the State Bar to comply with the data request. Mike McKee, *Calif. Supreme Court Rejects Professor's Bar Data Research Effort*, THE RECORDER, Sept. 23, 2008, http://www.law.com/jsp/ca/PubArticleFriendlyCA.jsp?id=1202424720703. Notwithstanding the concerns of anonymity expressed above, the legal argument for denying access to the data by the State Bar was that the Bar was not legally obligated to do so under either California common law or the more recent Proposition 59, which protects the "public's right to attend open court proceedings and to review documents that reflect those proceedings and adjudications made therein." *Id.* This petition was denied on September 17, 2008, without prejudice to re-filing in an appropriate court. *Id.*

A petition was then submitted to the Superior Court, and it was denied on March 29, 2010. Sander v. State Bar of California, CPF 08-508880, Proposed Statement of Decision (Mar. 24, 2010), *viewable at* http://www.box.net/shared/gf9laj5f20. The Court held that the purpose for the request, i.e., to examine the mismatch hypothesis, was irrelevant to the request for the data. *Id.* Rather, the crux of the matter was whether state common law and Proposition 59 could be applied. Here, the Court explained that Sander had not provided a principled argument for obtaining the data, and that such a request, if granted, would imply that all information held by a public agency should be made available upon request. *Id.* The Court concluded that "The law applicable to the courts before Proposition 59 was not that broad; and there is no evidence that the proposition was intended to work such a radical change." *Id.* Sander and his colleagues filed an objection to this decision on April 7, 2010, arguing, among other things, that "Proposition 59 creates a qualified right of access to records not expressly exempt from disclosure under California constitutional or statutory provisions, and that disclosure is required if there is no compelling justification for secrecy." Sander v. State Bar of California, CPF 08-508880, Petitioner's Objections to Proposed Statement of Decision (Apr. 7, 2010), *available at* (http://www.seaphe.org/pdf/petitionersobjections.pdf).

39. The discussion later in this article of existing research highlights the salience of this threshold issue.

40. Other interesting research, such as estimating mismatch effects, might be carried out by law schools themselves. The effects of changes in admissions policies on law school outcomes could also be investigated.

41. If anonymity were the central issue, there is a technical solution for disguising bar passage rates for both individual students and law schools. The randomized response method was devised precisely for such a purpose. For example, suppose the goal is to disguise a response for a particular person regarding whether she or he passed the bar exam. The response can be disguised by a data collection agency as follows. For each person in a data set, a (virtual) coin is flipped. The response is coded to "yes"

it seems unlikely that the State Bar of California would willingly provide randomly disguised data (for bar outcome and race/ethnicity), and it also seems unlikely that researchers from Project SEAPHE would be satisfied by a data file with racial and ethnic identifiers scrubbed.

To flesh out issues of why racial data and analyses may or may not be important, we consider in more detail two very different motivations for further research in this area. First, if the primary goal is to investigate mismatch, then disaggregation of the results by racial and ethnic status is not required. In fact, it is a harmful distraction. Clearly, no one has contended that the effects of mismatch are attributable to race or ethnicity in and of themselves. Rather, mismatch purportedly results from any admission preference that is inconsistent with a student's level of credentialing, e.g., a legacy admission in law school, or an athletic admission in undergraduate school. Second, even if race and ethnicity are taken as rough proxies for mismatch, it is nonetheless mismatch that is being investigated, rather than affirmative action. Yet this proxy approach is problematic since race and ethnicity are flawed indicators of (or instruments[42] for) affirmative action admission, and statistical comparisons between racial and ethnic categories do not produce a definitive evaluation of affirmative action—even though reasonable guesses might be made regarding the effects of measurement error.[43] Moreover, as we show below it is unlikely that simple regression modeling can compensate for this error in a uniform manner across racial and ethnic classifications. (And it is

---

if the coin comes up tails, and the actual response, if the coin comes up heads. Only the data agency knows whether the answer of an individual reflects the toss of the coin or actual outcome, and responses aggregated to a school level likewise contain some degree of distortion. Suppose the overall proportion passing the bar examination is *p*. The coin flip divides the passers into two randomly equivalent groups (heads and tails). Thus, for the group that flipped tails and responses truthfully, the expected value of the proportion of passers is *p*/2. Doubling the observed proportion then gives the desired proportion passing. It should be added that randomized responses can be included as outcomes in statistical models, but the technique so far has been primarily used to estimate incidence for dichotomous variables, such as a graduation or pass rate. *See* ARIJIT CHAUDHURI & RAHUL MUKERJEE, RANDOMIZED RESPONSE: THEORY AND TECHNIQUES (PIN) (2008). In practice, however, technical solutions are highly challenging given the limitations of the data available. For example, data from bar exam applicants often exclude students who did not graduate; attrition rates vary greatly by school, and school sample size often includes a very small number of affirmative action admits. This creates significant problems if the unit of analysis is at the school level.

42. Some researchers have used a Black-White indicator variable as an instrument in the estimation approach known as *instrumental variables* for obtaining the effect of affirmative action. *See, e.g.*, Jesse Rothstein & Albert H. Yoon, *Affirmative Action in Law School Admissions: What Do Racial Preferences Do?*, 75 U. CHI. L. REV. 649 [hereinafter "*Affirmative Action"*].

43. *See* Jesse Rothstein & Albert H. Yoon, *Mismatch in Law School* (Nat'l Bureau of Econ. Research, Working Paper No. 14275, 2008), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=881110 [hereinafter "*Mismatch in Law School*"].

worth stressing that in no argument we have encountered has anyone attributed the mismatch effect to actual racial or ethnic characteristics.) The potential outcomes model, discussed in the next section, provides a better methodological framework.

A distinction here should be drawn between policy making and research. Evaluators and researchers may be willing to disaggregate (rather than to compare) any "mismatch effects" by race and ethnicity for the purpose of investigating the general effects of affirmative action. Yet as noted above, the use of racial and ethnic identifiers as proxies most appropriately serves to identify mismatch effects rather than racial effects, regardless of whether mismatch effects are suspected *a priori* to be larger in some groups than others. For the purpose of policy making, there is no reason why particular demographic categories should be singled out. It could certainly be argued that research should use this proxy strategy, especially because race and ethnicity are generally available as variables in student databases whereas the degree of individual mismatch is not. But policy makers can be easily confused into thinking that the results contain direct lessons about affirmative action. So while researchers may sensibly continue to use the race and ethnicity fields in databases to explore possible mismatch effects, a more principled course of action would be to ignore race and ethnicity in the search for match effects—a course of action dependent on obtaining more direct measures of mismatch.

Finally, if the intention is to argue against admission preferences on the basis of negative match effects, then the symmetric position is to argue in favor of admission preferences in particular cases where positive match effects are encountered. If, for instance, researchers find that lower-credentialed students gain a relative advantage when admitted into elite law schools,[44] would professor Heriot and other proponents of meritocracy in law school admissions welcome this symmetry into the logic of their arguments? If not, this suggests a fundamental incompatibility between affirmative action perspectives based on efficiency or meritocracy and those based on mismatch.

IV. METHODOLOGICAL CONSIDERATIONS

At the center of the mismatch hypothesis is a counterfactual: students admitted to institutions where their academic credentials are below the average *would have learned more at a less elite law school*. Consequently, such students are placed at risk for graduation, bar passage, and ultimately joining the ranks of the profession as practicing lawyers. This phenomenon, if it is in fact happening, would certainly be troubling as a matter of policy and perhaps also as a matter of law. This core counterfactual helps to frame the literature synthesis and the statistical

44. The research discussed later in this article concerning K–12 grouping practices would suggest that this might very well be the case.

methods section of this article. Specifically, the counterfactual approach has been formalized in the potential outcomes model,[45] which has been adopted by most of researchers in the social sciences (e.g., economics, medicine, and education) who investigate match effects in both law school and undergraduate admissions.[46]

Richard Sander has explicitly and implicitly framed the mismatch hypothesis counterfactual is several related ways. In *Systemic Analysis*, he asked, "What would have happened to minorities receiving preferences had the preferences not existed?"[47] Other variations, in turn, have different implications for admission practices. For example, "If one is at risk of not doing well academically at a particular school, one is better off attending a less elite school and getting decent grades."[48] Another variation is found in his *Reply to Critics*:

> A large number of those receiving large preferences will struggle academically, receive low grades, and actually learn less in some important sense than they would have at another school where their credentials were closer to the school median. The low grades will lower their graduation rates, bar passage rates, and prospects in the job market.[49]

These questions (and assertions) are all directly or implicitly counterfactuals, and they exemplify the historical development of counterfactual reasoning as an important framework for understanding causation in the empirical social sciences, especially with respect to controversial issues.[50]

---

45. For an introduction, see Paul W. Holland, *Statistics and Causal Inference*, 81 J. AM. STAT. ASSOC., 945 (1986).

46. A number of recent papers use the potential outcomes models, albeit differently, to examine mismatch in law school and undergraduate education. *See, e.g.*, Sigal Alon & Marta Tienda, *Assessing the "Mismatch" Hypothesis: Differences in College Graduation Rates by Institutional Selectivity*, 78 Soc. Educ. 294 (2005); *Mismatch in Law School*, *supra* note 43.

47. *Systemic Analysis*, *supra* note 3, at 368.

48. *Id.* at 445. Three additional variations on the counterfactual in that article are as follows: "The principal question of interest is whether affirmative action in law school generates benefits to blacks that substantially exceeds the cost to blacks." *Id.* at 369; "The principal 'cost' I focus on is the lower performance that usually results from preferential admissions . . . If the struggling leads to lower grades and less learning, then a variety of bad outcomes may result. . . ." *Id.* at 370. "In a less competitive school, the same student might well thrive because the pace would be slower, the theoretical nuances would be a little less involved, and the student would stay on top of the material. The student would thus perform better in an absolute as well as a relative sense." *Id.* at 450.

49. *Reply to Critics*, *supra* note 3, at 1966.

50. Holland attributed this approach to Donald Rubin. Holland, *supra* note 45, at 946. *See also* Donald B. Rubin, *Estimating Causal Effects of Treatments in Randomized and Non-randomized Studies*, 66 J. OF EDUC. PSYCHOL., 688 (PIN) (1974). Others have traced aspects of this model to R.A. Fisher, J. Neyman, A.D. Roy, and L.L. Thurstone. *See* Jerzy Splawa-Neyman, *On the Application of Probability Theory to*

The usefulness of counterfactual modeling has been debated by historians but has become broadly accepted in the social sciences.[51] It has a particularly appealing (methodologically speaking) application in studying law school admissions. This is in part because of the strong parallel between Sander's formulation of the mismatch as a hypothetical question and the formal logic of the counterfactual model. Once a student enters law school, the proposed match effect is the net gain or loss relative to her hypothetical performance had she attended a less elite institution—which is similar in interpretation to a value-added effect. A simple (or naïve, as statisticians say) comparison of average outcomes across elite and non-elite schools is distorted by differences in incoming credentials. More formally stated, the naïve comparison is subject to selection bias resulting from the fact that students at more elite schools generally have higher incoming credentials. This, in turn, creates a tougher pool within which to compete for grades but also creates a pool more likely to achieve success

Agricultural Experiments. Essay on Principles, Section 9, 5 STAT. SCI. 465 (Dorota M. Dabrowska & Terence P. Speed trans.) (1990); A.D. Roy, *Some Thoughts on the Distribution of Earnings*, 3 OXFORD ECON. PAPERS, 135 (1951); L.L. Thurstone, *A Law of Comparative Judgement*, 34 PSYCHOL. REV., 278 (1927). The introductory material provided in this article is not intended to be taken as a comprehensive treatment. A well-known use of the counterfactual method in historical research methods is the 1964 work *Railroads and American Economic Growth* by Robert Fogel, in which he invented the counterfactual method to understanding the economic impact of railroads on the U.S. economy. *See* ROBERT W. FOGEL, RAILROADS AND AMERICAN ECONOMIC GROWTH: ESSAYS IN ECONOMETRIC HISTORY (1964). The Royal Swedish Academy cited this study in awarding Fogel the Nobel Prize for economics in 1993. *The Prize in Economics 1993 - Presentation Speech*, NOBELPRIZE.ORG, http://nobelprize.org/nobel_prizes/economics/laureates/1993/presentation-speech.html (last visited May 24, 2011). Prior to this development, the dramatic U.S. economic growth from about 1865 to 1890 had been popularly attributed to the expansion of the railroad system. The book details Fogel's approach to use quantitative methods to create a counterfactual world in which the U.S. canal and road systems were developed as alternatives to rail transportation. Based on this quantitative construction, he determined the level of per capita income achieved at the beginning of 1890 would have been achieved only three months later had railroads not been built.

51. Counterfactual reasoning in history has been disparaged by historian E. H. Carr as a "parlor game." EDWARD HALLETT CARR, WHAT IS HISTORY? 127 (1961). Others have advanced a more nuanced opinion. For example, M. Bunzl characterized counterfactual reasoning as coming in "two varieties—good and bad. The bad reasoning is bad because it has no grounding; it is merely an act of imagination, and unconstrained imagination at that. The good reasoning is good because it can be grounded." Martin Bunzl, Counterfactual History: A User's Guide, 109 Am. Hist. Rev., 845 (2004). About the time of Fogel's work, simultaneous developments of the counterfactual approach to determining causality were also occurring in a wide variety of disciplines including philosophy, economics, and statistics. *See*, *e.g.*, DAVID K. LEWIS, COUNTERFACTUALS (2001); James Heckman, *Shadow Prices, Market Wages, and Labor Supply*, 42 ECONOMETRICA, 679; Richard E. Quandt, *The Estimation of the Parameters of a Linear Regression System Obeying Two Separate Regimes*, 53 J. AM. STAT. ASSOC., 873; Richard E. Quandt, *A New Approach to Estimating Switching Regressions*, 67 J. AM. STAT. ASSOC., 306.

on the bar exam and in employment.[52]

In the statistical sciences, the counterfactual framework is known as the "potential outcomes model." Briefly, in any situation where there is a focus or treatment group (T) and a reference or control group (C), there are two potentially different outcomes for each individual depending on whether she receives the treatment or is part of the control group. Here, attending an elite law school is akin to the treatment condition, and attending a non- or less-elite school is akin to the control condition. Let the outcomes for these two cases be denoted as Y^T for T and Y^C for C, so the treatment effect can be expressed as

$$\text{Equation (1): } \Delta = Y^T - Y^C$$

The effect of interest, if the counterfactual is premised on preferential admissions, is then defined as the average Δ for students given preferential admission to elite law schools. The fundamental problem with estimating this average effect is that for any individual, it is only possible to measure the effect under one condition—*either* the treatment (an elite law school) *or* the control condition (a non-elite law school), but not both. The basic idea of a matching analysis is that for each individual in group *T* a similar individual from group *C* is found based on a particular set of background factors denoted as *X*. A treatment effect for an individual *i* from group *T* *controlling* for *X* (which is abbreviated to $| X$) is then

$$\text{Equation (2): } \Delta_i = (Y^T_i - Y^C_i) \mid X$$
$$\approx (Y^T_i \mid X) - (Y^C_j \mid X)$$

Here the outcome for individual *j* matched on *X* is the counterfactual or *what if* outcome for individual *i* had she or he attended a non-elite school (note the change from *i* to *j* from the first line of Equation (2) to the expanded second line). The average of the difference Δ_i[53] is then taken over all *n* students attending elite law schools ($i = 1, 2, \ldots, n$), in order to arrive at the desired effect. In statistical literature, this effect is usually dubbed the *average treatment effect for the treated*, or *ATT*.[54]

52. The elite–non-elite distinction has limited precision. It is a practical rather than an ideal way of describing the effect of attending a school with more as opposed to less stringent admissions criteria.

53. Statisticians call this the *expected value* as opposed to the average effect.

54. Consider an example of the *ATT*. Students at parochial schools often appear to outperform students at public school. Using the counterfactual method, the question becomes "How would a parochial school student have performed had he or she attended a public school" (or vice versa). Some studies have shown that the apparent advantage is negligible from the standpoint of the *ATT* when family and other

This *ATT* appears to be closely related to the mismatch hypothesis as formulated in *Systemic Analysis*.[55] However, it is also the case that a counterfactual difference (in the potential outcomes model) according to the mismatch hypothesis should vary by the degree of mismatch. That is, a larger negative effect should be evident for students who are relatively less credentialed than their peers. A number of researchers have formulated models with this feature.[56] It is important to add that comparisons between members of different racial or ethnic groups, made using the potential outcomes model, cannot provide an estimate of a causal effect of being a member of such a group. This is because a student cannot be randomly assigned to such categories. As explained by measurement expert Paul Holland, formerly of the Educational Testing Service, "For causal inference, it is critical that each unit be potentially exposable to any one of the causes. As an example, the schooling a student receives can be a cause, in our sense, of the student's performance on a test, whereas the student's race or gender cannot."[57] This is important because the gap in outcomes

---

background factors are controlled. In fact, it has been shown that the achievement of students who attended parochial schools might have been higher had they attended public schools. *See*, *e.g.*, STEPHEN L. MORGAN & CHRISTOPHER WINSHIP, COUNTERFACTUALS AND CAUSAL INFERENCE: METHODS AND PRINCIPLES FOR SOCIAL RESEARCH (2007) (providing more detail on the effectiveness of parochial schools); Sarah Theule Lubienski & Christopher Lubienski, *School Sector and Academic Achievement: A Multi-Level Analysis of NAEP Mathematics Data*, 43 AM. EDUC. RES. J. 651 (2006). *See also* Stephen L. Morgan, Counterfactuals, Causal Effect Heterogeneity, and the Catholic School Effect on Learning, 74 SOC. EDUC., 341 (2001).

55. *Systemic Analysis*, *supra* note 3, at 429. Yet two versions of Sander's work display the same conflation as Heriot's reasoning. Accordingly, the mismatch hypothesis is stated both as a counterfactual and as a purely descriptive statement about the effect of student qualification on law school outcomes:

> "In other words, the collectively poor performance of black students at elite schools does not seem to be due to their being 'black' (or any other individual characteristic, like weaker educational background, that might be correlated with race). The poor performance seems to be simply a function of disparate entering credentials, which in turn is primarily a function of the law schools' use of heavy racial preferences."

As noted above, both versions cannot be simultaneously true. Sander provides an anecdote supposedly clarifying the mismatch hypothesis in which he recalls having performed poorly in an elementary language class at Harvard and speculated that he would have performed better in a class with less talented (in language capacity) peers. *Id.* at 449–50. However, if the course material were the same and were tested similarly, there is no reason to believe that Sander's counterfactual performance would have been better. Given these assumptions, one could as easily speculate that less talented peers might have reduced his expectations for achievement. *Id.*

56. *See*, *e.g.*, Rothstein & Yoon *supra* note 42, at 659-60; Douglas Williams, *A Review of the Econometric Literature on Law School Mismatch*, Paper Presented at the Annual Meeting of the Law and Society Association (May 25, 2009) (unpublished; abstract at http://www.allacademic.com/meta/p304241_index.html.

57. Paul W. Holland, *Statistics and Causal Inference*, 81 J. AM. STAT. ASSOC., 945, 946. Holland and Rubin invented the phrase "No causation without manipulation." PAUL W. HOLLAND, EDUCATIONAL TESTING SERVICE, CAUSATION AND

between Black and White students is central to some past analyses of the mismatch hypothesis. For instance, while one could ask whether the Black-White gap is more likely to be larger at elite than at non-elite schools,[58] Holland would argue that this is not a sensible *causal* question because it subsumes a racial comparison that cannot be manipulated.[59]

V. HETEROGENEOUS LEARNING ENVIRONMENTS

To date, the focus of the mismatch literature has been on the *effect* of mismatch on student outcomes (intermediate and longer-term) rather than the precipitating cognitive or social mechanisms. That is, a quantitative estimate can be obtained and might be thought of as an estimate of a *causal effect* in the potential outcomes framework, but this offers little illumination of the *mechanism leading to* any such effect.[60] As originally presented by Sander, the starting point for a causal mechanism was conceptualized in terms of a norm-referenced measure such as class rank, which embodies a kind of relative competitive pressure for grades.[61] This approach is flawed, however, because by definition there will always be a lower tail of the grade distribution. If there is substantial variation, as one would expect with any larger educational institution, then whoever is at the bottom will be mismatched and will learn less. In any classroom or law school with considerable heterogeneity, there will always be a group of students on the left tail of the distribution, whose "preparation and cognitive skills" are "substantially less developed" than most of their peers.[62] This approach makes the idea of a "mismatch" relatively meaningless.

RACE RESEARCH REPORT 03-03 (2003), *available at* http://www.ets.org/Media/Research/pdf/RR-03-03-Holland.pdf [hereinafter EDUCATIONAL TESTING]. In other words, "causes are experiences that units undergo and not attributes that they possess . . . ." *Id.* at 8. Holland argued that "[t]he useful role of RACE is its ability to reveal varying effects of interventions on different parts of a diverse population . . . ." *Id.* at 19.

58. EDUCATIONAL TESTING, *supra* note 57, at 3.

59. *Id.* Some economists nonetheless use the Black–White comparison as a key element of some statistical models. For example, see Rothstein & Yoon, *supra* note 42.

60. Of course, the lack of precision (or correctness) in the description of the mechanism does not invalidate the observation of a negative match effect. As noted by Paul W. Holland, "The description of a causal mechanism (How?) can be completely wrong while at the same time the effect of the cause (What if?) is clear and replicable." EDUCATIONAL TESTING, *supra* note 57, at 7.

61. Sander seems to support the use of class rank for such a purpose. "In other words, it was not the absolute ability of a student that determined staying power in the traditionally more difficult natural science majors, but rather the student's ability relative to his or her peers." *Systemic Analysis*, *supra* note 3, at 452.

62. *See* THE SCALE AND EFFECTS OF ADMISSIONS PREFERENCES IN HIGHER EDUCATION (PROJECT SEAPHE), THE EFFECT OF LAW SCHOOL RACIAL PREFERENCES ON MINORITY BAR PERFORMANCE, B-2 (2007), http://www.seaphe.org/pdf/bar-proposal/project_description_revised_proposal.pdf [hereinafter SCALE AND EFFECTS].

Moreover, this phenomenon is independent of affirmative action—or any admission policy that does not somehow ensure homogeneity on whichever characteristics result in mismatch. Based on the dataset that has been used for most existing studies of the mismatch hypothesis, only about fifty percent of students within any racial or ethnic category who are most likely to attend elite schools based on background features (credentials) actually attend elite schools. Conversely, about fifty percent of students who are least likely to attend elite schools do in fact attend elite schools.[63] Thus, there is a substantial amount of mixing in student credentials and backgrounds, even for White students within elite schools—mixing that would *not* diminish if affirmative action were discontinued.[64] Finally, it should be recognized that the homogeneity of entering credentials would have to remain fairly static during the students' three years at the law school; if students start growing or declining relative to one another, the mismatch will reappear. In short, if the problem is posed in normative terms, any solution linked to admissions will be very hard to carry out.

A number of scenarios have been offered about how mismatch in cognitive skills might translate into diminished student outcomes. Rothstein and Yoon provided one scenario about a Black student admitted to highly selective School X under affirmative action:

> There, she is a small fish in a big pond: Nearly all of her white classmates enter law school with stronger academic credentials, more experience with legal concepts, and stronger writing skills than she has. She works hard, but the academic demands at School X are much higher than at School Y, and by the end of the first year she finds herself near the bottom of her class. She does not make law review, and will graduate—if she does—without academic distinction.[65]

Another plausible explanation of the mismatch mechanism was given by Williams:

> The mismatch hypothesis begins with the assumption that classroom instruction is pitched to the median student. If this is the case, students too far below the median may struggle to

63. *See* Gregory Camilli & Darrell D. Jackson with Chia-Yi Chiu & Ann Gallagher, *The Mismatch Hypothesis in Law School Admissions*, 2 WIDENER J. LAW, ECON. & RACE 165, 204 (2011).

64. To some degree, this results from the variability of schools within tiers in the database and is a weakness of the crude classification of law schools into six categories (as discussed later in this article). However, it seems unlikely that the preponderance of this mixing is wholly due to measurement error inherent in the classification scheme, because the clustering procedure used to create that dataset shows a substantial degree of separation between elite and non-elite schools. *See* LINDA F. WIGHTMAN, LAW SCHOOL ADMISSIONS COUNSEL, CLUSTERING U.S. LAW SCHOOLS USING VARIABLES THAT DESCRIBE SIZE, COST, SELECTIVITY, AND STUDENT BODY CHARACTERISTICS, Research Report 93-04 (1993).

65. *Affirmative Action*, *supra* note 42, at 659-70.

> understand class discussions and to keep up with the pace of instruction. Consequently, mismatched students learn less and may even reduce their effort if they become discouraged, leading to even less human capital accumulation.[66]

A third account is from Sander, who proposes a cognitive mechanism to go along with his relative (norm-based) explanation:

> 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 college freshmen trying to skip first-year physics to go straight into advanced classes: some students may learn faster, but many may crash and burn.[67]

Setting aside problems with the physics analogy (the content of first-year courses is pretty standard across law schools), this story-telling by various authors seems somewhat plausible. Yet there is virtually no evidence available that such processes are actually happening or are grounded in learning theory. Moreover, even if some students do suffer learning detriments, others may respond differently and benefit. A good hunch is no substitute for empirical evidence, and there are three useful sources of information—in the research literature concerning how students perform in heterogeneous learning environments—that do shed light on the causal mechanism. The first concerns studies of promotion and tracking in the literature in K–12 education, the second concerns student-school match in undergraduate education, and the third concerns the effect of admission preferences in law school.

A. K–12 Studies

"Mismatch" ideas have been explored for students as early as kindergarten, focused at that age on the idea of "readiness." This issue sometimes is framed as the idea of parental redshirting—waiting until children are six years old to enroll them in kindergarten—and is also sometimes framed around the possibility of grade retention between kindergarten and first grade.[68] The focus of this research is generally whether a younger student benefits from waiting and becoming an older

66. Williams, *supra* note 56, at 9.

67. SCALE AND EFFECTS, *supra* note 62, at B-2.

68. *See, e.g.*, Guanglei Hong & Stephen W. Raudenbush, *Effects of Kindergarten Retention Policy on Children's Cognitive Growth in Reading and Mathematics*, 27 EDUC. EVALUATION & POL'Y ANALYSIS 205 (2005); Guanglei Hong & Bing Yu, *Effects of Kindergarten Retention on Children's Social-Emotional Development: An Application of Propensity Score Method to Multivariate Multi-Level Data*, 44 DEVELOPMENTAL PSYCHIATRY 407 (2008); Lorrie A. Shepard & Mary Lee Smith, *Synthesis of research on school readiness and kindergarten retention*, 44 EDUC. LEADERSHIP, 78 (1986).

first-grade student, and whether an underperforming student benefits from retention that results in an additional year in kindergarten.[69] Focusing just on this latter issue, here is how the question can be asked in the counterfactual sense: "Would a retained student have performed lower in first or ensuing grades had that student been promoted?" For those who support such retention policies, a lower level of preparation in kindergarten means that a student will have an academic or behavioral mismatch to her first-grade peers. However, in a review by Lorrie Shepard of sixteen controlled studies on kindergarten retention, she found no academic or social benefits for students who had spent an extra year in kindergarten.[70] She also reported that "schools that do not practice kindergarten retention have just as high average achievement as those that do but tend to provide more individualized instruction within normal grade placements."[71]

The mismatch idea has also been explored for later grades, in the context of ability grouping, or 'tracking,' a practice that has been authoritatively denounced as harming students placed in lower tracks.[72] A National Research Council report recently recommended "that both formal and informal tracking by ability be eliminated. Alternative strategies should be used to ensure appropriately challenging instruction for students who vary widely in their skill levels."[73] Such detracking at the K–12 level has been shown to have great potential to increase both equity and overall outcomes.[74] The learning experiences underlying these results—harms of tracking and benefits of detracking—have been extensively researched, and three primary causal explanations have emerged: (a) stratified distribution of resources, including the most effective teachers; (b) peer effects; and (c) expectations effects.[75]

The issue of academic expectations seems particularly salient here, since it challenges the basic presumption of the mismatch hypothesis. That is, the research on tracking and expectations supports the conclusion that a

---

69. *Id.*

70. See Lorrie A. Shepard, *Negative policies for dealing with diversity: When does assessment and diagnosis turn into sorting and segregation?*, in LITERACY FOR A DIVERSE SOCIETY: PERSPECTIVES, PRACTICES, AND POLICIES, 279 (Elfrieda H. Hiebert ed., 1991).

71. *Id.* at 287.

72. *See* HIGH STAKES: TESTING FOR TRACKING, PROMOTION, AND GRADUATION (Jay Philip Heubert & Robert Mason Hauser eds., 1999).

73. NAT'L RESEARCH COUNCIL, ENGAGING SCHOOLS: FOSTERING HIGH SCHOOL STUDENTS' MOTIVATION TO LEARN 219 (2004).

74. *See* CAROL CORBETT BURRIS & DELIA T. GARRITY, DETRACKING FOR EXCELLENCE AND EQUITY (2008).

75. *See, e.g.*, JEANNIE OAKES, KEEPING TRACK: HOW SCHOOLS STRUCTURE INEQUALITY (2d ed. 2005); KEVIN GRANT WELNER, LEGAL RIGHTS, LOCAL WRONGS: WHEN COMMUNITY CONTROL COLLIDES WITH EDUCATIONAL EQUITY (2001); Carol Corbett Burris, Edward W. Wiley, Kevin G. Welner, & John Murphy, *Accountability, Rigor, and Detracking: Achievement Effects of Embracing a Challenging Curriculum as a Universal Good for All Students*, 110 TEACHERS COLL. RECORD 571 (2005).

more homogeneous, non-"mismatched" educational setting backfires; in lower-tracked classes, it results in what former President Bush has called the "soft bigotry of low expectations."[76] Students are insufficiently challenged and set on a path of academic failure.[77] The apparent assumption underlying the mismatch hypothesis—that a more homogeneous setting will result in greater learning for those who would otherwise be among the less qualified in the more elite learning environment—has also long been proposed by those supporting K–12 tracking systems, but the empirical results consistently show otherwise.[78] Accordingly, any finding of mismatch effects in law school (of "lower achievers" suffering a detriment as a result of being placed with a "higher achieving" group) would be inconsistent with research from grades one through twelve.[79]

B. Undergraduate Studies

Empirical research has also explored issues of mismatch in undergraduate education. Using regression analysis, Fischer and Massey[80] focused on Black and Hispanic undergraduates in a sample of students attending elite colleges and universities, and they examined three outcomes: GPA, leaving school, and perception of college and university success. They found small *positive* student-school match effects (representing student-school as the difference between a student's SAT score and the average SAT at a college or university). That is, a "mismatch" was associated with higher, not lower, performance.

Using the same data set with a broader sample of students, however, Massey and Mooney[81] found no student-school match effects for retention (staying in college or university) or hours studied. For a group of legacy-admitted students, and looking at *GPA* outcomes, they found a small negative effect in which mismatch in the group of legacy students was

76. George W. Bush, President of the United States, President's remarks in Minneapolis, Minnesota (October 30, 2004), http://georgewbush-whitehouse.archives.gov/news/releases/2004/10/20041030-8.html (last visited May 20, 2011).

77. *See* OAKES, *supra* note 72.

78. *See also* Carol Corbett Burris, Kevin G. Welner, & Jennifer W. Bezoza, Educ. and the Pub. Interest. Ctr. & Educ. Policy Research Unit, *Legislation Policy Brief: Universal Access to a Quality Education: Research and Recommendations for the Elimination of Curricular Stratification* (2009) http://nepc.colorado.edu/publication/universal-access (last visited May 20, 2011).

79. *See* Burris, Wiley, Welner, & Murphy, *supra* note 72.

80. Mary J. Fischer & Douglas S. Massey, *The Effects of Affirmative Action in Higher Education*, 38 SOC. SCI. RES. 531 (2007). Ordinary least squares was used for quantitative outcomes, and logistic regression for binary outcomes.

81. Douglas S. Massey & Margarita Mooney, *The Effects of America's Three Affirmative Action Programs on Academic Performance*, 54 SOC. PROBS. 99 (2007).

associated with lower performance for this group.[82] However, this result is greatly obscured by the lack of equivalence in *GPA* across schools.

Brand and Halaby[83] examined the effect of elite college or university attendance (i.e., the treatment in the counterfactual model) for mismatched students and for all students. They found that attending an elite school yielded occupational status benefits for the former group, but not for the latter group.[84] That is, they found no mismatch effect.

Alon and Tienda[85] estimated the elite–non-elite effect for Asian, Black, Hispanic and White students on a six-year graduation rate. Using an econometric modeling approach, they found that all students tended to benefit from attending elite schools.[86] Specifically, most estimated match effects were significantly positive with a few near zero, depending on grouping and modeling variations.[87] In sum, no support for the mismatch hypothesis was found.[88]

In a widely cited and influential study, Dale and Krueger compared life outcomes for students who were accepted and rejected by comparable schools. That is, they had identical admission decisions across sets of schools.[89] Because some of these students eventually attended more (and some less) elite schools, the researchers argued that this methodology controls for variables that may be observable to admission committees but not statisticians.[90] Though not the equivalent of random assignment, since unobserved variables likely played a role in the students' subsequent decisions to accept or reject an elite school's offer, the data allow nonetheless for a nice quasi-experiment.[91] Looking first at the general population of students, the researchers found no effect on life outcomes for increasing eliteness.[92] Of particular importance to those considering the mismatch hypothesis, Dale and Krueger also concluded "there is no

82. *Id.* at 113.

83. Jennie E. Brand & Charles N. Halaby, *Regression and Matching Estimates of the Effects of Elite College Attendance on Educational and Career Achievement*, 35 SOC. SCI. RES. 749 (2006).

84. *Id.* at 753. The average treatment effect or ATE combines two counterfactuals: elite-school students who might have attended non-elite schools (the ATT), and non-elite-school students who might have attended elite schools (the ATC, or average treatment effect for the untreated). Brand and Halaby's results suggest that a benefit would be obtained if students who attended non-elite schools had instead attended elite schools. This is the meaning of the ATC.

85. Alon & Tienda, *supra* note 46, at 302.

86. *Id.* at 306.

87. *Id.* at 307.

88. *Id.* at 306.

89. *See, e.g.*, Stacy Berg Dale & Alan B. Krueger, *Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables*, 117 Q. J. ECON. 1491 (2002).

90. *Id.* at 1492–93.

91. *Id.* at 1493.

92. *Id.* at 1492.

evidence in these data that students who score relatively low on the SAT exam do worse in the labor market by attending schools with a relatively high average SAT."[93] Looking specifically at the small sample of Black students in the dataset, they also concluded that Black students benefited (in the counterfactual sense) from attending elite schools just as much as other students in terms of subsequent earnings.[94]

C. Law School Studies

Sander's *Systemic Analysis* was based in large part on data from the *Bar Passage Study*[95] and, as of the time of its publication in 2004, constituted the most extensive investigation of the mismatch hypothesis in law school. Using a weighted index of LSAT and undergraduate GPA (UGPA),[96] Sander found that Black applicants to law school had the same probability of admission as White applicants with substantially higher academic index values.[97] He argued that LSAT and UGPA are the strongest predictors of

93. Stacy Berg Dale & Alan B. Krueger, *Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables*, National Bureau Of Economic Research Working Paper 7322, 23, http://www.nber.org/papers/w7322. *See also College Selectivity*, PROJECT SEAPHE, http://www.seaphe.org/topic-pages/college-selectivity.php (last visited May 24, 2011).

94. Dale & Kruger, *supra* note 85, at 1493. Dale and Krueger reported a small negative match effect for GPA, but they noted that GPA is not comparable across schools (in fact, a comparable problem exists regarding grades when studying law schools).

95. WIGHTMAN, *supra* note 64. To preserve the confidentiality of the data in the Bar Passage Study, the identity of individual law schools was omitted in the public-use data set. Instead, to allow other researchers to study the relationship between school characteristics and student outcomes, the law schools were empirically clustered into six groups based on median Law School Admission Test (LSAT) score, median undergraduate grade point average (UGPA), tuition and fees, enrollment, selectivity, percent minority and faculty/student ratio. The six clusters were described as follows: 1) Elite, 2) Public Ivy, 3) Second tier public, 4) Second tier private, 5) Third tier, and 6) Historically Black (note that these categories, particularly the Historically Black category, are not set forth as an ordinal ranking). According to Wightman, 24,814 (about 60% of the entering cohort) consented to the release of their law school bar exam performance record, and 93% of these students had graduated and taken the bar exam during the course of the 6-year study. Evidence-based arguments regarding the mismatch hypothesis and affirmative action in law school admissions continue to be driven by data that are now over one decade old.

96. *Systemic Analysis*, *supra* note 3, at 381. Sander used the formula for academic index (*AI*): $AI = 0.6 * LSAT + 0.4 * UGPA$ linearly rescaled to the range 0–1000. Using a weighted combination of LSAT and UGPA scaled to the interval [0, 1000], Sander found that that Black applicants to law school had the same probability of admission as White students with academic index values of about 140 points higher. *Id.*

97. *Id.* at 431, tbl. 5.3. Note that it is the absolute amount of achievement (knowledge and skill tested), perhaps compared to others in a given state, not relative achievement within a given law school, that should affect performance on the bar examination. Accordingly, relative standing in a law school class in terms of LGPA does not equate to the amount of learning, achievement, or (directly) bar exam success. Indeed, even students in the lowest decile have most likely learned a great deal upon

first-year Law School GPA (LGPA)—much stronger than law school eliteness or race/ethnicity.[98] Because law school grades (or at least relative grades within any given law school) showed the strongest relationship to bar passage, he argued that the strong downward pressure on Black students' grades due to preferential admission at more elite schools eventually translates into lower bar passage rates.[99]

The model implicit in the original version of the negative match hypothesis offered in *Systemic Analysis* thus takes a rudimentary form that does not directly include bar passage rates as a measured outcome. Rather, the argument was that achievement (as measured by law school grades[100]) has a stronger effect on bar passage than the combination of incoming credentials and increased academic proficiency resulting from elite school attendance. Consequently, any factor that lowers LGPA is presumed to lower the probability of passing the bar examination.[101]

Yet, as Ho pointed out,[102] researchers interested in investigating the mismatch hypothesis should be engaging with a more complete question. Even assuming a tightly linked relationship between LGPA and bar passage rates, the direct learning benefits of attendance at higher-tier schools may compensate, to some degree, for any negative effect on bar passage due to the downward pressure (in elite schools) on LGPA resulting from mismatch. Just looking at those two factors (which is still a simplified model), the total effect of admission preference is then conceptualized as the sum of the two causal chains.[103] In any event, the methodological sophistication of mismatch studies has evolved substantially from this modest beginning, and a range of statistical models and nonparametric[104] matching techniques have been used more recently by both proponents and skeptics of the mismatch hypothesis.[105] A number of these are discussed briefly below; however, these approaches tend to be mathematically complex, and detailed descriptions are outside the scope of this article.

Ayres and Brooks, also looking at law schools, used a strategy roughly

---

graduation.

98. *Id.* at 439, Tbl 5.6. *See also id.* at 444, tbl. 6.1.

99. *Id.* at 479.

100. *Id.* at 411. It is not clear whether achievement is taken to mean the absolute or relative level of skill and ability.

101. *Id*. at 422–23.

102. Daniel E. Ho, *Why Affirmative Action Does Not Cause Black Students to Fail the Bar*, 114 YALE L.J. 1997 (2005). *See also* Daniel E. Ho, *Affirmative Action's Affirmative Actions: A Response to Sander*, 114 YALE L.J. 2011 (2005).

103. Rothstein & Yoon in *Affirmative Action*, *supra* note 42, at 3-4 (formalizing this argument).

104. In nonparametric techniques, weaker assumptions are made about the relationship between analytic variables. For example, linearity is not assumed.

105. Sander has more recently proposed a "case control" methodology that bears a strong resemblance to propensity score matching. SCALE AND EFFECTS, *supra* note 62, at B-1, B-10.

similar to that which Dale and Krueger used for undergraduate education.[106] They found mixed support for the negative match (mismatch) hypothesis. Two groups of students were identified: those who attended their first-choice school and those who were accepted at their first choice school but chose to attend a presumably lower-tier school.[107] Because both groups of students were selected by their first-choice schools, this helps to control for unobserved variables that contribute to student success, and thus to decrease selection bias. Based on the comparison between first-choicers and alternative-choicers, they concluded that first-year grades[108] and first-attempt bar passage outcomes appeared to lend marginal support for the mismatch hypothesis.[109] However, they observed non-significant differences when examining outcomes having farther reaching significance[110] (e.g., graduation rates and ultimate bar passage).[111]

Rothstein and Yoon proposed a methodology in which parameters for estimating match effects are constructed with two different statistical models.[112] In the first, Black students at elite and non-elite schools are compared, and in the second, Black students are compared to White students—assuming race is a proxy for affirmative action preference.[113] The models were chosen to roughly provide upper and lower bounds for the match effect.[114] In other words, the two statistical procedures were chosen in order to sandwich an unbiased statistical estimate.[115] Both

106. Ian Ayres & Richard R. W. Brooks, *Does Affirmative Action Reduce the Number of Black Lawyers?*, 57 STAN. L. REV. 1807 (2005). Ayres and Brooks also carried out an analysis to determine the effects of eliminating affirmative action on the flow of Black lawyers into the profession, so-called stereotype threat as a possible explanation of Black underperformance in law school, and the potential policy of advising candidates more fully based on their probability of success in law school. We do not address this question herein though we are in full agreement that these are important topics.

107. *Id.* at 1832.

108. *Id.* That is, comparable students tended to get higher grades at non-elite schools.

109. *Id.* at 1835.

110. *Id.* at 1838. That is, comparable students tended to graduate at similar rates, whether they attended elite or non-elite schools.

111. In *Reply to Critics*, *supra* note 3, Sander appeared to have embraced the methodology of Ayres and Brooks. Upon comparing first-choice Black students with Black students who passed up their first-choice school, he estimated the mismatch effect to be -14.9% for first bar examination and -3.6% for ultimate bar passage. *Id.* at 1994, tbl. 7. Sander reported that he was not able to replicate the results of Ayres and Brooks. *See* Rothstein and Yoon, *supra* note 42, at 681-82 (further critique of the Ayres and Brooks methodology).

112. *See* Mismatch in Law School, supra note 43.

113. *Id*, at 2.

114. *Id.*

115. *Id.* In particular, Rothstein and Yoon used two types of models: ordinary least squares (OLS) and instrumental variables (IV). The IV approach is popular among econometricians, but in the current context requires the identification of a variable that is correlated with an outcome variable (e.g., bar passage) only through its effect on

models obtained match estimates for graduation and bar passage that were not significantly different from zero.[116] But with both models the effects for post-graduation employment outcomes were positive.[117] That is, mismatched students in elite schools had better employment outcomes than had they been "matched" to non-elite schools.[118]

Rothstein and Yoon contend that their sandwich strategy increased confidence in the disconfirmation of the mismatch hypothesis.[119] However, their results hinged on a significant methodological choice: the lowest quintile of all students on the academic index (composed of LSAT and UGPA) was eliminated from the analysis, which in turn eliminated most (about seventy-five percent) of all Black students in the sample from their data analysis.[120] The researchers did not include these students because, they argued, the majority of White students with low levels of qualification are normally excluded even from the least selective law schools.[121] That is, even the least selective schools would normally not admit a particularly low-scoring White student, so those White students who are in fact admitted are likely higher on unobserved variables than other White students with the same qualification; thus, any Black–White comparison would be biased in favor of White students.[122]

Williams provided a "distance" framework for understanding mismatch similar to the approach used by Fischer and Massey, and by Massey and Mooney.[123] He reviewed the methodology of Rothstein and Yoon, of

---

selection to elite versus non-elite schools. Rothstein and Yoon in a technical argument showed how race could be used as an instrument with resulting statistical bias in one direction, while the statistical bias for their OLS estimate of the match effect would be in the other direction.

116. *Id.* at 19.

117. *Id.*

118. *Id.* at 22.

119. *Id.*

120. *Id.* at 17–18.

121. *Id.* at 18.

122. Rothstein and Yoon argue that "even if black and white applicants would have achieved similar average outcomes, there is reason to expect that those white students who actually matriculated would have outperformed the average black applicant even in the absence of affirmative action." *See Affirmative Action*, *supra* 42 at 696. Note that a White student with similar credentials at a non-elite school is used here as the counterfactual for a Black student attending an elite school. While we recognize that some economists have proposed the use of such counterfactuals within the context of a thoughtful model, the comparison is tantamount to making the case that a Black student is the same as a White student for the purpose of estimating an outcome, given that suitable control variables can be identified—or that the direction of the bias created by unobservable variables can be guessed. We think it will be difficult to establish a broadly appealing argument with this strategy.

123. Williams, *supra* note 56. In brief, "distance" for an individual is based on the difference between that individual's academic index and the average or median index at a particular school. Distance then serves as an independent variable in evaluating outcomes.

Barnes, and of Ayres and Brooks, which were all conducted with the *Bar Passage Study* dataset.[124] Williams then carried out analyses roughly similar to those of Rothstein and Yoon for law school graduation and ultimate bar passage, but he also constructed a new outcome measure for bar passage that gave greater weight in the analysis to students who passed the bar with fewer attempts.[125] The argument was that such a measure is more closely aligned to learning.[126] A number of statistically significant and negative match effects were found, though only after omitting students from the middle two tiers of law schools (second-tier public and second-tier private) from the analysis.[127] That is, the counterfactual to elite law school attendance was that the student would attend a very non-elite school. Williams argued that eliminating the middle tiers would reduce measurement error in the classification of law schools.[128]

Even more so than with the Rothstein and Yoon study, however, this approach places clear emphasis on a methodological choice that may affect external validity. Eliminating those "second-tier" categories removes from the analysis the most convincing counterfactual students, and thus decreases the quality of the ATT estimator. It also raises the question of whether this comparison has many real-world (as opposed to modeled) examples. Students attending UCLA tend to be substantially different from those attending Podunk State. The comparison only to lower-tier law schools also raises a related methodological question: whether the study is comparing applicants so substantially different that it is beyond the capacity of parametric regression models to control for those differences. Though the intent of Williams' analysis is clear, elimination of a substantial proportion of a sample in order to produce an effect is clearly open to further discussion. Indeed, Sander criticized Rothstein and Yoon[129] using exactly this argument.[130]

In any case, the study's model yielded estimates suggesting that Black students at elite law schools pass the bar at a rate of eight to twelve percentage points lower than similar students at non-elite schools on their first attempt, and a rate of five to ten percentage points lower on their final attempt.[131] Analyses similar to that of Ayres and Brooks were also carried out, with the result being no effect for ultimate (as opposed to first attempt)

---

124. *Id.*
125. *Id.*
126. *Id.*
127. *Id.*
128. *Id.*
129. *Affirmative Action*, *supra* note 42.
130. *See* Richard H. Sander, *Are Black/White Disparities in Graduation and the Bar Getting Better, or Worse?*, EMPIRICAL LEGAL STUDIES BLOG (Sept. 19, 2006, 8:28 AM), http://www.elsblog.org/the_empirical_legal_studi/2006/09/page/3/.
131. *Id.* Williams also controlled for regional difference in bar exam difficulty.

bar passage.[132] Williams concluded that there is substantial reason to believe that evidence exists in the *Bar Passage Study* data to support the negative mismatch hypothesis.[133]

A study by Camilli and his colleagues is a more recent examination of the mismatch hypothesis.[134] Like the others, they used the *Bar Passage Study* data.[135] They carried out separate analyses for all racial and ethnic groups, using bar passage rates as the outcome of interest.[136] Students with the same *a priori* chances of being admitted to an elite school (based on twelve admission qualifications and background factors) were considered to be comparable.[137] For any group of students with similar chances within a racial or ethnic category, the researchers were able to identify some

---

132. *Id.*

133. *Id.*

134. *See* Camilli et al., *supra* note 63.

135. *Id.*

136. *Id.*

137. *Id.* After controlling for these twelve variables, a sensitivity analysis suggested the estimate was robust with respect to other variables available in the *Bar Passage* data. Specifically, neither SES nor geographic region (in which the bar examination was taken) had a notable effect on match estimates.

It should be noted that Camilli and his colleagues imputed missing data, while in the Williams study no mention is made of how missing data were treated. *See* Camilli et. al., *supra* note 63. *See also* Williams, *supra* note 56. This is a key concern—how researchers using the *Bar Passage Study* have compensated for missing information in the dataset. Missing data were explicitly taken into account in the Camilli et al. study, in contrast to virtually all previous studies in which missing data procedures were not mentioned. The improper practice of deleting cases with missing values may potentially bias both model coefficients and their standard errors. The problems with missing data are 1) loss of efficiency, 2) complication in data handling and analysis, and 3) bias due to unknown systemic trends in the unobserved data. It is well known that mean substitution or pairwise deletion does not account for the variation that would be present if the variables are observed, resulting in downward bias in the estimation of variances and standard errors. *See also* RODERICK J. A. LITTLE & DONALD B. RUBIN, STATISTICAL ANALYSIS WITH MISSING DATA (2d ed. 2002); Phillip L. Roth, *Introduction to the Feature on Problematic Data*, 6 ORG. RES. METHODS 279 (2003). Mean substitution and pairwise deletion methods "provide problematic estimates in almost all instances." As implemented multivariable normality is assumed, this method has been widely used for imputing missing values for binary variables such as bar passage, but this procedure does not compensate for 3) above. *See also Therese* D. Pigott, *A Review of Methods for Missing Data*, 7 EDUC. RES. AND EVAL. 353 (2001). Because binary variables are treated like normal variables in the imputation steps, imputed values may typically be fractional.

One strategy for imputing the binary data is to round up or down the imputed fraction to 1 or 0. However, it has been shown that such rounding can produce substantial bias, and it is generally recommended to use the unrounded imputed values for analysis. *See also* Nicholas J. Horton, Stuart R. Lipsitz, & Michael Parzen, *A Potential for Bias When Rounding in Multiple Imputation*, 57 AM. STAT. 229 (2003). *See also* Christopher F. Ake, *Rounding After Multiple Imputation With Non-binary Categorical Covariates*, SUGI 30 PROCEEDINGS, (2005) http://www2.sas.com/proceedings/sugi30/112-30.pdf.

students who attended elite schools and some students who attended less elite schools.[138] The latter were used as counterfactuals for the former.[139] When controlling for the propensity to attend an elite school, Camilli and colleagues found non-significant but negative effects of -5% and -1.4% for the first and ultimate bar passage rates (respectively) for Black students who attempted the bar examination.[140] In contrast, Sander in his *Reply to Critics* reported estimates of difference in bar passage rate of -14.9% and -3.6%, respectively.

In a finer-grain breakdown of match effects, Camilli and colleagues also subdivided each racial and ethnic group into two data sets: one with higher *a priori* chances of attending law school (i.e., higher propensities) and one with lower chances.[141] Roughly speaking, this provides one solution to Williams' concern that match effects have a bias toward zero resulting from the imprecise classification of schools into tiers. Lower-chance students are much more likely to have attended a lower-tier school, but they can be matched to elite-schools students of the same racial or ethnic group. It could be expected that the elite–non-elite comparison for lower-chance students would magnify negative effects, due to a greater degree of mismatch. The results did in fact suggest a possible negative match effect for lower-chance Black students; the match effects were insignificant but negative: -7.7% and -5.6% for first and ultimate bar passage, respectively.[142] For higher-chance Black students, the corresponding effects were again insignificant but the magnitudes were mixed: -2.9% and +3.4%, respectively.[143] These findings, though not statistically significant, might be suggestive of negative match effects for some students with low *a priori* probabilities of being admitted to an elite school, although not for Black students with stronger credentials. The estimates of -7.7% and -5.6% would probably apply to at most about forty percent of Black student applicants, and this figure may be even lower.[144] Rothstein and Yoon similarly showed that bar passage rates fell sharply only for Black students in the lowest twenty percent of the Black distribution on the admissions index who attended elite schools.[145] These results suggest that those concerned about negative match effects should be focusing only on least-

138. *Id.*

139. This is known as propensity score matching. *See* Daniel E. Ho, Kosuke Imai, Gary King, & Elizabeth A. Stuart, *MatchIt: Nonparametric Preprocessing for Parametric Causal Inference*, 42 J. STAT'AL SOFTWARE 481 (2011). *See also* Jasjeet S. Sekhon, *Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R*, 42 J. STAT'AL SOFTWARE 1 (2011).

140. *Id.*

141. *Id.*

142. *Id.*

143. *Id.*

144. *Id.*

145. *Affirmative Action*, *supra* note 42, at 691.

credentialed applicants.

Camilli and colleagues also found some (statistically insignificant) evidence in support of the mismatch hypothesis for Asian law school students, although not for Latino students, in the lower-chance range.[146] *Negative* match effects were obtained for White students in the lower-chance range; though these were small, they were statistically significant.[147] Yet when looking at higher-chance students (who generally have higher levels of qualification) in all racial/ethnic categories, the match effects for bar passage were generally close to zero or positive—though none approached statistical significance at $\alpha = .05$. When breaking down results by gender, only two notable effects were observed. A positive, but non-significant match effect of 7.5 percentage points was found for Black women on first-time bar passage (adjusted for number of attempts).[148] And a strong positive match effect of thirteen percentage points was found for lower-chance Hispanic men on adjusted first-time bar passage; this effect was statistically significant at $\alpha = .05$.[149]

Given that these relatively low-credentialed Hispanic males are clearly among the intended beneficiaries of affirmative action policies, this last finding seems particularly important. We are not arguing that this finding should be used to support affirmative action policies aimed at such students. Rather, this finding suggests the weakness of global arguments against affirmative action based on the *Bar Passage* data. At most, such findings may provide signals of cultural differences, which in turn may provide clues for enhancing student learning. As Roxana Moreno put it:

> Expert teachers never assume that a particular student will think or behave in a manner that is expected for his/her gender, culture, or SES but rather view each student as a unique individual and use what they know about group differences to help explain why students learn differently in school.[150]

D. Summary

We began this discussion of evidence regarding heterogeneity and learning by suggesting the research literature might shed light on the causal mechanism driving any mismatch effects. However, there is no compelling case of the existence of such a mechanism prior to law school. In fact, the literature suggests a reverse mechanism, with greater challenge leading to an increase in achievement—a positive "mismatch" effect. But there may be no mystery here. As the above discussion of mismatch research regarding law schools shows, the existing research base fails to document a

146. *Id.*
147. *Id.*
148. *Id.*
149. *Id.*
150. Roxana Moreno, EDUCATIONAL PSYCHOLOGY 27 (2010).

consistent and substantial negative mismatch effect at that level either. Some studies suggest positive effects, some suggest negative effects, and some suggest no significant effects. If enough snark hunters return empty handed, there is not much reason to examine or explain the nature of snarks.[151] Though there is a suggestion of negative effects for some Black students, these effects do not consistently rise to the level of statistical significance; indeed, the significance levels *within* Williams' study vary according to methodological choices.

Nonetheless, it is worth exploring interpretations and implications of any negative match effects. The research from K–12 and undergraduate college and university education suggests, as we discuss in the conclusions below, that any negative match effects observed in law school are more likely due to the practices of law schools or unobserved (unmeasured) characteristics of students rather than the credentials. If this is the case, then the phrase "mismatch hypothesis" embodies a misspecification of the causal mechanism. Any effect might better go by the humdrum label "inadequate support hypothesis." Research exploring effective learning environments and successful practice would then follow. Other research might examine how success in law school should be measured in terms of the intended curriculum. For example, it might be useful to investigate empirically the claim that elite law schools tend to focus more on esoteric and national issues, while less elite schools tend to focus more on the content of the state bar examination. While the research conducted to date does suggest the possibility of a small to moderate negative match effect for some (e.g., relatively low-credentialed) students, it should be recognized (a) that methodological choices might account for whether a given match effect is or is not statistically significant,[152] and (b) that a significant match effect signals, but does not identify, a causal mechanism.

VI. RETHINKING BENEFITS, OUTCOMES, MEASURES, AND IMPLICATIONS

The benefits of attending an elite law school are not easily captured. Chief Justice Vinson explained this more than sixty years ago in *Sweatt v. Painter*, a case finding that separate law school for Black applicants was not "equal" to the law school at the University of Texas:

> What is more important, the University of Texas Law School

151. The origin of snark hunting is found in LEWIS CARROLL, THE HUNTING OF THE SNARK: AN AGONY IN EIGHT FITS (1876). The creature in Carroll's story may not in fact exist. *See also* Medellín v. Texas, 552 U.S. 491, 549 (2008) (Breyer, J. dissenting, finding the majority's insistence on finding some indication of self-executing intent in a treaty's text to be akin to "hunting the snark").

152. Ayres and Brooks made a similar point: "What are the underlying institutional factors that undermine black law students' chances of becoming lawyers, as compared to white law students with the same entering credentials attending the same tier schools? No responsible educator can ignore this question or fail to take action." Ayres & Brooks, *supra* note 97, at 1854.

> possesses to a far greater degree those qualities which are incapable of objective measurement but which make for greatness in a law school. Such qualities, to name but a few, include reputation of the faculty, experience of the administration, position and influence of the alumni, standing in the community, traditions and prestige. It is difficult to believe that one who had a free choice between these law schools would consider the question close.[153]

Correspondingly, it is difficult to capture the benefits of diversity on the law schools themselves. As the Supreme Court articulated in *Grutter*, a diverse student body "promotes learning outcomes and better prepares students for an increasingly diverse workforce, for society, and for the legal profession."[154]

Whereas the individual merit model implies a few narrow criteria (such as law schools' grades or graduation rates) in operationalizing "greater benefit to society," the reasoning in *Grutter* implies that characteristics of a student body are intrinsic to high-quality legal education. Chief Justice Vinson's opinion in *Sweatt* touches on this as well:

> The law school, the proving ground for legal learning and practice, cannot be effective in isolation from the individuals and institutions with which the law interacts. Few students and no one who has practiced law would choose to study in an academic vacuum, removed from the interplay of ideas and the exchange of views with which the law is concerned.[155]

Similarly, the individual merit model fails to capture everything about future lawyers that might be of importance. A recent report of prediction of effectiveness by Shultz and Zedeck attempted to identify non-cognitive factors relevant to the practice of law, such as situational judgment or past attitudes and experiences.[156] They constructed new non-cognitive predictors based on biographical information and social judgment as well as ratings of lawyering effectiveness on twenty-six different dimensions (e.g., research and information gathering; conflict resolution; and entrepreneurship).[157] They found low to moderate correlations between these new predictors and most of the effectiveness scales. Based on these results, they observed:

> The impressive aspect of these results was (1) the large number of Effectiveness Factors that were predicted by the [biographical

153. Sweatt v. Painter, 339 U.S. 629, 634 (1950).

154. *Grutter*, 539 U.S. at 308.

155. *Painter*, 339 U.S. at 634.

156. MARJORIE M. SHULTZ & SHELDON ZEDECK, FINAL REPORT: IDENTIFICATION, DEVELOPMENT AND VALIDATION OF PREDICTORS FOR SUCCESSFUL LAWYERING (2008), *available at* http://ssrn.com/abstract=1353554.

157. *Id.* at 4.

> information] and the [social judgment] tests, and (2) that the correlations were generally higher, though moderately so, than those between the LSAT and the small subset of Effectiveness Factors that overlap with the LSAT and with which it had an expected relationship (e.g. Analysis and Reasoning, Researching the Law, Writing).[158]

A reasonable conclusion from this and similar work is that no single predictor or measurement instrument will come close to capturing everything important about a future lawyer. As noted by Sackett and Lievens, "With cognitively loaded predictors as generally the strongest correlates of task performance and noncognitive predictors as generally the best predictors in the citizenship and counterproductive behavior domain, careful attention to the criterion of interest to the organization is a critical determinant of the eventual makeup and success of a selection system."[159] While it is not clear that non-cognitive factors can currently be incorporated into admission criteria, development of tools in this area is ongoing.[160]

From this perspective, the fundamental issue is that narrow admissions criteria are at best loosely coupled with an array of unobserved non-cognitive factors that lead to effective lawyering. In fact, one benefit of attending a more elite school consists of access to professional networks and organizations, within which these non-cognitive qualities are important determinants of success. The potential for public leadership and private practice is partially determined by academic preparation, but whether this is the lion's share of success is an open question. Relatively exclusive reliance on standard admission criteria is more a function of their measurability; this is more a pragmatic choice than a utility or merit maximizing argument.

The key measure used in mismatch hypothesis research—bar passage rates—is also somewhat problematic. While passing a state bar exam is a very important outcome, it does not necessarily map equally well onto the curriculum of different law schools. If it is true that less elite law schools teach relatively more state bar content while more elite schools teach relatively more national law and abstract theory, then the bar examination does not accurately measure the different learning of students in the different types of schools. If an applicant opts for a more elite law school, she may be sacrificing bar preparation instruction for other instruction of value. According, the lesser-qualified student who attends an elite school might have done better at a different institution not because he or she learned more, but rather because the other institution taught more of the bar

158. *Id.* at 80.

159. Paul R. Sackett & Filip Lievens, *Personnel Selection*, *in* 59 ANN. REV. PSYCHOL. 419, 422 (Susan T. Fiske et al. eds., 2008).

160. *Id.*

material. Measurement expert Al Beaton once remarked, "If you want to measure change, don't change the measure."[161] In the context of this paragraph, we would offer the alternative, "If you want to measure game, don't game the measure."

Finally, as noted earlier in this article, the existing mismatch hypothesis research is troublingly tied to the debate about affirmative action and therefore troublingly assumed to have primary policy implications related to law school admissions policies. But a clear finding of the research about K–12 ability grouping is that heterogeneous learning environments are most successful when supports are provided for teachers and students. Even setting aside that research, it stands to reason that if incoming credentials are highly correlated to first-year success in law schools, then students preferentially admitted will need additional supports. To the extent that we do find a mismatch effect for the lowest-credential students, a reasonable conclusion is that law schools need to do a better job in providing such learning supports. Any healthy discussion of a mismatch effect should involve programmatic interventions by which negative match effects can be addressed. No one, after all, has claimed that the unobserved characteristics to which mismatch effects might be attributed are endowed or unalterable—or that the unobservables relate in any way to merit.

VII. CONCLUSION

As a policy matter, the primary push behind the mismatch hypothesis is to question the benefits of affirmative-action admissions policies. A key issue, therefore, is whether current support for the mismatch hypothesis is strong enough to support a change in admission policies, and there are several important considerations. First, research regarding mismatch in law school is mixed: some positive estimates and some negative estimates have been obtained, and results vary by methods of statistical design and analysis, outcome analyzed, race, and gender. A reasonable conclusion is that current data do not support a robust finding of statistically significant mismatch effects. The most pointed conclusion that current analyses would support is only that there appear to be some mismatch effects when looking at first-time bar passage rates for the least credentialed applicants, although even those effects are small to moderate and are inconsistent.

Even if one were to conclude that negative match effects exist for, e.g., the least credentialed Black students, the mismatch hypotheses put forward informally by economists and others are logically suspect. No causal

161. *See* George W. Bohrnstedt, *U.S. Mathematics and Science Achievement: How Are We Doing?*, 99 TCHRS C. REC. 19, 22 (1997). Similarly, a quantitative comparison cannot be validly made if the measure changed. This is known as the instrumentation threat to internal validity. *See also* WILLIAM R. SHADISH, THOMAS D. COOK, & DONALD T. CAMPBELL, EXPERIMENTAL AND QUASI-EXPERIMENTAL DESIGNS FOR GENERALIZED CAUSAL INFERENCE (2002).

mechanism has been offered to explain the phenomenon with roots either in cognitive theory or the extant empirical research on heterogeneous learning environments. In contrast to mere intuitive hunches about how student-school mismatch hinders learning, hard evidence is available that a greater degree of challenge in heterogeneous learning environments benefits students. Moreover, if the mismatch mechanism is unrelated to race or ethnicity, then observed mismatch effects *must* be due to unobserved background characteristics that are correlated with race or ethnicity and not controlled by LSAT or GPA. From this perspective, match effects, to the degree they exist, eventually arise from an unobserved capacity for development that is present at the moment of enrollment. Moreover, proponents of the mismatch hypothesis appear to be arguing that, independently of observed qualifications, admission under an affirmative action policy is a proxy for these unobserved variables—with race then being a proxy for such affirmative action admission. However, we know of no empirical work identifying the cognitive mechanisms. Neither competitive pressure as indicated by class rank nor intuitive notions self-evaluation (possibly including discrimination or stereotype threat) are compelling explanations. To the contrary, the broader research literature on heterogeneous learning environments suggests that higher expectations may result in *positive* effects for mismatched students.

In any case, the conceptualization of a match effect as due to unobserved student characteristics may lead to a more fruitful search of influences on student learning. For example, in one scenario, a match effect arises indirectly as an interaction between what is *not* taught and the experience of incoming students; that is, some students may not have been previously (prior to admission) exposed to material in the *intended* curriculum through their college or university preparation, extra-curricular activities, or informal learning. This lack of pre-exposure could lead to learning difficulties which could manifest as a negative match effect relative to students with sufficient exposure, given a substantial gap between the intended curriculum and what is actually taught, i.e., the *received* curriculum. Also consider the scenario in which two students, who are indistinguishable in terms of measured credentials, are offered admission to the same schools (of which some are more and some less elite). These students would appear to be highly comparable in terms of academic potential, given that admission committees considered both observed characteristics of those students as well as other qualities ascertained from their applications. (Note that the latter qualities, perhaps set forth in their admissions essays, are typically not available to secondary analysts and in this sense are unobserved). Now suppose one student chooses an elite school based on a preference for status and the other student chooses a less elite school after conducting financial projections and evaluating social supports. This latter student has planned in a mature, thoughtful way, which may indicate qualities leading to more successful law school

outcomes, and could lead to relatively higher performance at the less elite school. This phenomenon could also manifest as a negative match effect. In both of these scenarios, identification of the unobserved variables might have useful implications for pre-law guidance, or academic preparation or support in law school.

Despite these methodological concerns, the technical evaluation of the mismatch hypothesis is just one framing issue. Even if match effects were estimated with a much higher degree of precision, there is only a tenuous connection between match effects and the logic in *Grutter v. Bollinger* concerning the broad educational benefits of diversity. Empirical match effects have no bearing on the intangible benefits of attending and graduating from an elite law school, nor do they have bearing on the "merit" argument against affirmative action. In the end, the two issues—match and affirmative action—may inform one another, but they are truly separate conversations. Evidence of mismatch ultimately has nothing to do with race or ethnicity, and explanatory support for the hypothesis should be examined with regard to students substantially less credentialed, whatever the basis of their admission. Focusing on race, and thereby on affirmative action, sorts out one group and thus creates the perception of a political motivation to change admission policies. This may further encourage the incorrect assumption that without affirmative action, there would be little diversity in pre-law credentials at elite schools. Most unfortunately, the affirmative action focus of the extant mismatch hypothesis discussion implies an admissions-based solution to any performance concern, rather than a solution grounded in academic and social supports for lower-achieving students.