

Reply

Affirmative Action's Affirmative Actions: A Reply to Sander

Daniel E. Ho

I am grateful to Professor Sander for his interest in my work and his willingness to pursue a valid answer to the critical question of the effects of law school tier on bar performance. Sander's readiness to respond to my Comment¹ demonstrates the importance of the questions at hand and his openness to progress on these issues. Fortunately, progress is possible, because, as I show here, the impressive-sounding points in Sander's Response² violate basic methodological principles and are incorrect.

Sander points to certain descriptive facts that my Comment does not dispute. Black students appear to fail the bar at higher rates than white students. It also appears that "blacks and whites with similar *law school grades* (when controlling for school and entering credentials) have virtually identical graduation and bar outcomes."³ However, these descriptive observations are irrelevant to the causal question of whether going to a higher-tier law school causes black students to fail the bar. As my Comment and this Reply demonstrate, black law students who are similarly qualified when applying to law school perform equally well on the bar irrespective of what tier school they attend. There is no evidence that affirmative action reduces the bar performance of the students it is designed to help. The descriptive facts Sander presents may account for some of the *reasons* for affirmative action, but they do not address the *consequences* of affirmative

* All analyses presented in this Reply are available at <http://www.yalelawjournal.org>.

1. Daniel E. Ho, Scholarship Comment, *Why Affirmative Action Does Not Cause Black Students To Fail the Bar*, 114 YALE L.J. 1997 (2005).

2. Richard H. Sander, *Mismeasuring the Mismatch: A Response to Ho*, 114 YALE L.J. 2005 (2005).

3. *Id.* at 2006 (emphasis added).

action. Here, I respond to each of Sander's points in turn.

First, Sander's control group, as he conceives it, is invalid. As I noted in my Comment, every law school examined practices some form of affirmative action, so the data is not informative about the broad effects of affirmative action. Instead, we may investigate the effects of attending a higher-tier law school. In response, Sander asserts that white students are the control group and black students are the treatment group, because only blacks were admitted under a system of affirmative action.⁴ This conception is incorrect, because black students could differ from white students in all sorts of respects that are not due to affirmative action or law school tier. This flaw is akin to assigning estrogen to a group of women and a placebo to a group of men, and inferring the effect of estrogen on health outcomes by comparing women and men. The black/white test score gap has been documented across a host of fields, so other differences (like primary school education, family structure, or culture) may be at work.⁵ Comparing blacks and whites in order to infer the causal effect of law school tier ignores all of these other differences between these groups and thereby provides an invalid estimate of the causal effect.⁶

Second, Sander asserts that controlling for law school grades does not bias the estimate of the causal effect of law school tier. He does not take issue with the rule of inference that controlling for a consequence of the cause is never justified and will never produce the right causal effect.⁷ This rule is undisputedly violated here, because Sander's analysis itself demonstrates that going to a higher-tier law school decreases law school grades⁸ and that law school grades are correlated with bar passage.⁹ Sander's responses to this basic flaw do not withstand scrutiny. Standardizing grades within schools is no solution, because the same student may receive lower grades in a higher-tier school even if she learned the same amount. Moreover, the fact that law school GPA is the variable with the biggest effect in Sander's regressions¹⁰ is not surprising at all. In fact, it is an indicator of the problem: This coefficient reflects the fact that his regression is controlling away the consequences of law school tier, the

4. *Id.*

5. See THE BLACK-WHITE TEST SCORE GAP (Christopher Jencks & Meredith Phillips eds., 1998).

6. Sander himself recognizes this problem, finding it necessary to argue that "being black" is not an "independently significant causal factor" determining bar passage outcomes. Richard H. Sander, *A Systemic Analysis of Affirmative Action in American Law Schools*, 57 STAN. L. REV. 367, 447 (2004). The coefficient on being black is only insignificant in the original bar passage regression, however, when Sander (incorrectly) controls for law school grades. *Id.* at 444.

7. See GARY KING ET AL., *DESIGNING SOCIAL INQUIRY: SCIENTIFIC INFERENCE IN QUALITATIVE RESEARCH* 78 (1994); PAUL R. ROSENBAUM, *OBSERVATIONAL STUDIES* 73-74 (2d ed. 2002).

8. Sander, *supra* note 6, at 373.

9. *Id.* at 443.

10. See Sander, *supra* note 6, at 443; see also Sander, *supra* note 2, at 2007.

key causal variable. As my Comment demonstrates, omitting law school grades from the original regressions leads the coefficient on law school tier to be indistinguishable from zero and the coefficient on black students to be substantially negative. This result shows that within Sander's analysis there is no detectable effect of law school tier and that preexisting credentials alone do not account for the black/white bar passage gap.

To further investigate this issue, Sander now proposes a structural equations model.¹¹ This represents a different approach based on different assumptions from the original article, whereas the whole point of my reanalysis was to *reduce* the role of unfounded and unnecessary assumptions. Sander's discussion fails to do precisely what my Comment aimed to achieve, namely, to clarify the assumptions in substantively meaningful ways in order to validly infer a causal effect. Clarifying the assumptions of Sander's structural equations model shows that it in fact *supports* each of my Comment's main points: (1) that we cannot control for law school grades in estimating the causal effect of tier, (2) that white students are an invalid control group, and (3) that there is no detectable effect of law school tier on bar passage. To verify this, note that the path diagram itself readily admits post-treatment bias by the arrows from tier to law school GPA to bar passage. As a result, the total effect in this path diagram simply represents the coefficient of tier in a bar passage regression,¹² (correctly) excluding law school GPA. In addition, the model eliminates white students, making it inconsistent with Sander's (invalid) "organiz[ation] around a comparison of 'treatment' blacks . . . and 'control' whites."¹³ Most importantly, applying this exact structural model to the original data set and variables yields a tier effect that is indistinguishable from zero.¹⁴

If this model supports all of the points in my Comment, why has it been introduced? Not only has Sander's Response changed the assumptions and the model, but it has also changed the outcome variable from eventual bar passage, as defined in his original article, to first-time bar passage.¹⁵ Applying this structural model to the original outcome variable of *eventual*

11. *Id.* at 2007-08.

12. See Paul W. Holland, *Causal Inference, Path Analysis, and Recursive Structural Equations Models*, 18 SOC. METHODOLOGY 449, 457 (1988).

13. Sander, *supra* note 2, at 2006.

14. This is estimated with a linear probability model of bar passage controlling for LSAT score and undergraduate GPA, excluding gender, because that is what is depicted in Sander's path diagram. Using a logistic regression or including gender does not alter the fact that the tier coefficient is insignificant in the original data set. Although Sander has not provided sufficient information, he appears to have applied ordinary least squares for a dichotomous outcome, which is wrong in part because predictions are not even bounded to the unit interval.

15. Although Sander describes the outcome as "whether a person passes the bar on one of her first two attempts," Sander, *supra* note 6, at 444, the user's guide for the data indicates that this variable actually represents eventual bar result, see LINDA F. WIGHTMAN, USER'S GUIDE: LSAC NATIONAL LONGITUDINAL DATA FILE 15 (1999).

bar passage, we find no tier effect. With the recoded outcome, the model does lead to a borderline significant effect of tier, but not at the significance level of .01 that Sander reports for all other coefficients.¹⁶ The marginal effect of going to a higher-tier law school in this model is roughly a 2.4% increase in the probability of first-time bar passage, plus or minus 2.5% at a 99% confidence level.¹⁷ This effect is neither statistically distinguishable from zero nor close to explaining the substantial black/white bar passage gaps as Sander originally claimed. (Recall that bar passage gaps are as large as 21% in the index range defined by Sander.)¹⁸ And of course if one conducts enough tests, with enough specifications, measures, models, and recoding, one can induce a statistically significant result even if the relationship is random (classic “Type 1 error”). Indeed, this problem of “model dependence” is precisely what matching methods are designed to remedy. Using exact matching for Sander’s newly defined outcome, thereby eliminating unjustified assumptions, the conclusion remains the same: Tier has no detectable effect on first-time bar passage for students who are similar in all other respects.

Third, Sander asserts that my analysis “ignores law school performance.”¹⁹ My analysis excludes law school grades precisely *because* it investigates the causal effect of law school tier on bar passage. I do not dispute Sander’s finding that going to a higher-tier law school reduces grades. Such a finding isn’t surprising—after all, we would expect that a higher-tier law school presents a more competitive environment. Moreover, I do not dispute that law school grades are correlated with bar passage. Precisely for these reasons, we cannot control for law school performance to assess the causal effect of law school tier. This is a textbook example of bias induced by controlling for a consequence of the cause.

Sander also makes two specific criticisms of my analysis, each of which applies equally to his original analysis. First, Sander claims that law school tier is mismeasured and hence that the results of my analysis do not prove much.²⁰ Yet if law school tier, the key causal variable, is mismeasured, the original analysis fails as well. Sander cannot have it both ways: The assumptions required to interpret the matching estimates causally are uniformly less onerous than those for his regression analysis. Asserting that the assumptions are violated proves the point that Sander’s original inference that law school tier causes black students to fail the bar cannot

16. The fact that all individual coefficients in this path diagram are statistically significant at .01 does not mean that the total effect of tier is.

17. This substantive effect is glossed over by standardization, which does not accord with a substantive quantity of interest and does not make the estimates comparable. See Gary King, *How Not To Lie with Statistics: Avoiding Common Mistakes in Quantitative Political Science*, 30 AM. J. POL. SCI. 666, 669-74 (1986).

18. Sander, *supra* note 6, at 446.

19. Sander, *supra* note 2, at 2009.

20. *Id.*

be sustained.²¹

Second, Sander argues that there are unobservable differences between similarly qualified students that invalidate matching.²² As I stated in my Comment, unobservable differences in these students would invalidate both my reanalysis and Sander's analysis.²³ Both matching and regression assume that there are no remaining unobserved differences in students of different tiers when holding constant pretreatment variables.²⁴ Unless the researcher gathers more data (as in my accompanying paper,²⁵ which controls for 170 additional covariates), we have no idea what impact unobservables might have on the estimates. Sander's claim that matching maximizes bias along unobservables compared to regression is a conjecture without any theoretical basis. The claim depends on the correlation between unobserved and observed variables, and as such it cannot be tested in the data.²⁶ For example, if higher-tier students have a higher (unobserved) underlying propensity to take the bar in "tough" jurisdictions like California or New York, the true causal effect of law school tier could actually be *positive* if we do not control for region.²⁷ In short, Sander's charge of unobserved variables concedes that the estimates of his own original analysis are unfounded.

If matched students cannot be viewed as anywhere close to randomly assigned to any reliable measure of law school tier,²⁸ then we cannot put credence in the estimates of the causal effect of law school tier presented by Sander or my Comment. The purpose of outlining the hypothetical experiment is to lay bare such critical assumptions for drawing a causal inference, thereby permitting scholars to assess the credibility of the estimates, which is what my Comment aimed to achieve.

21. Sander asserts that matching "singles out those pairs of students for whom 'tier' is least meaningful and most biased." *Id.* at 2010. If students differ in unobservable respects, however, we cannot attribute differences in bar passage rates to law school tier at all.

22. *Id.*

23. Ho, *supra* note 1, at 2001.

24. To be clear, some regression techniques account for some types of unobserved heterogeneity, but not the logistic regression employed by Sander.

25. Daniel E. Ho, Evaluating Affirmative Action in American Law Schools: Does Attending a Better Law School Cause Black Students To Fail the Bar? (Mar. 9, 2005) (unpublished manuscript), available at <http://www.yalelawjournal.org> and <http://people.iq.harvard.edu/~dho/research/sander.pdf>.

26. The standard advice in the literature is to first obtain balance on observables and then to conduct sensitivity analyses to examine to what degree inferences would change if some unobserved variable were correlated to the treatment and the outcome. In such an analysis, inferences could certainly change, but the actual truth cannot be ascertained from the data, because the data is by definition observed. See Paul R. Rosenbaum & Donald B. Rubin, *Assessing Sensitivity to an Unobserved Binary Covariate in an Observational Study with Binary Outcome*, 45 J. ROYAL STAT. SOC'Y SERIES B (METHODOLOGICAL) 212 (1983).

27. Cf. P.J. Bickel et al., *Sex Bias in Graduate Admissions: Data from Berkeley*, 187 SCIENCE 398 (1975) (reversing a finding of gender discrimination in graduate school admissions by controlling for academic department).

28. See Sander, *supra* note 2, at 2010.

To conclude, I want to underscore one basic point of agreement with Sander: The empirical investigation of affirmative action is important and should be subjected to scientific scrutiny. The fortunate fact is that we are not alone in this venture. Tools for analysis have been developed across academic fields, providing some easy fixes to reassess the mismatch hypothesis with more credible and theoretically consistent assumptions. This reanalysis shows that similarly qualified black students perform just as well on the bar irrespective of law school tier. Of course, law school tier may affect many other outcomes, like public service, life satisfaction, and career success, each of which merits independent investigation. This data set thereby provides a great opportunity for legal scholars to capitalize on developments in other academic disciplines, helping us to better understand the effect of law school tier and possibly of affirmative action.