Response of James P. Scanlan to Office of Management and Budget Request for Information "Methods and Leading Practices for Advancing Equity and Support for Underserved Communities Through Government" (FR Doc No: 2021-09109)

Responder: James P. Scanlan, Attorney at Law, Washington, DC, Tel: (202) 337-3927, Email: jps@jpscanlan.com

Areas of the RFI the response addresses: The assessment and methods aspects of areas 1 to 4.

Responder Expertise: I am an attorney who has spent a long career in and out of government in dealing with civil rights matters, while usually specializing in the use of statistics to measure racial and other demographic differences. I was the lead counsel for the U.S. Equal Employment Opportunity Commission (EEOC) in *EEOC v. Sears, Roebuck and Co.,* 504 F. Supp. 241 (N.D. Ill. 1986), affirmed, 839 F.2d 302 (7th Cir. 1988), which was likely the largest discrimination case ever fully tried and which was based almost entirely on statistical evidence. On leaving the EEOC in 1995, I was the agency's Assistant General Counsel for Expert Services responsible for overseeing the statistical cases prosecuted by agency's field offices.

I have written scores of articles on the measurement of demographic differences in the law and the social and medical science and have given methods workshops on that subject at some of the nation's leading universities. All recognitions of the pattern whereby the rarer an outcome the greater tends to the relative (percentage) difference between rates of experiencing the outcome and the smaller tends to be the relative difference between rates of avoiding the outcome, including by the National Center for Health Statistics, are based on my work. In the context of the instant RFI, the pattern may be described in terms that reducing any barrier to participation in a program, while tending to reduce relative demographic differences between rates of overcoming the barrier, will tend to increase relative demographic differences between rates of failure to overcome the barrier.

I gave been invited to testify before the U.S. Commission on Civil Rights (CCR) (Dec. 8, 2017) and meet with staff of the Department of Education (DOE) (Mar. 22, 2018) concerning school discipline disparities issues and to participate on an expert panel convened by the Department of Housing and Urban Development (HUD) (Sept. 22, 2020) concerning housing discrimination issues. My written testimony before the CCR and a handout I provided for the DOE meeting are attached as Attachments 1 and 2. A memorandum I provided to other participants on the HUD panel that addresses many of the issues I address here is attached at Attachment 3.

Introduction

This response addresses the assessment component of each of the first four subject areas listed in the RFI and explains the unsoundness of virtually all efforts at quantifying differences in the circumstances of advantaged and disadvantaged groups involving favorable or corresponding adverse outcomes as result of failure to understand how measures of differences between outcome rates tend to be affected by the prevalence of an outcome. It also explains how demographic differences involving outcome rates can be effectively quantified.

The principal focus of the response is the pattern whereby reducing barriers or otherwise increasing favorable outcomes and reducing the corresponding adverse outcomes, while tending to reduce relative racial and other demographic differences in rates of experiencing the favorable outcomes, tends to increase such differences in rates of experiencing the adverse outcome. An important focus is also the longstanding mistaken belief of governments agencies and others that reducing an adverse outcome will tend to reduce, rather than increase, relative differences in rates of experiencing the outcome.

The misunderstanding of this issue is part of a larger problem in the measurement of demographic differences arising from a general failure to understand how measures of differences involving favorable and corresponding adverse outcome rates tend to be affected by the prevalence of an outcome. That failure has caused researchers and policymakers to fail to address the crucial issue of the extent to which an observed pattern of change in a measure is solely a function of the change in the prevalence of an outcome and the extent to which it may reflect an actual change in the strength of the forces causing outcome rates of advantaged and disadvantaged groups to differ (including any such actual change that might result from a particular policy).

Discussions of varying comprehensiveness and complexity of both the larger problem and the issue that is the principal focus of this response may be found in my "Race and Mortality Revisited," *Society* (July/Aug. 2014),[1] Comments for the Commission on Evidence-Based Policymaking (Nov. 14, 2016) (CEP Comments), Letter to the American Statistical Association (Oct. 8, 2015) (regarding measurement issue pertaining to a wide range of subjects); Memorandum to HUD September 22, 2020 Expert Panel (Sept. 19, 2020, updated Jan. 15, 2021) (HUD Panel Mem.) (Att. 3) (regarding measurement issues pertaining to a range of matters then currently receiving great public attention); "The Mismeasure of Health Disparities," *Journal of Public Health Management and Practice* (July/Aug. 2016), Letter to Harvard University (Oct. 9, 2012), "Measuring Health and Healthcare Disparities," Federal Committee on Statistical Methodology 2013 Research Conference (regarding measurement issues pertaining mainly to health and healthcare disparities issues); amicus curiae brief in *Texas Department of Housing and Community Development, et al. v. The Inclusive Communities Project, Inc.*, Supreme Court No. 13-1731 (Nov. 17, 2014) (TDHCD Brief), "The Perverse Enforcement of Fair Lending

---

[1] I have not attempted any consistency of formatting for the various references. In some cases, references are identified solely by links. A copy of this document that may be subsequently corrected or annotated is available here. If it is corrected or annotated, such fact will stated on the cover and corrections or annotations will be identified.

Laws," *Mortgage Banking* (May 2014) (regarding measurement issues mainly pertaining to lending issues); "The Mismeasure of Discrimination," Faculty Workshop, University of Kansas School of Law (Sept. 20, 2013) (Kansas Law Paper) (pertaining to identification and quantification of discrimination); and "Measuring Discipline Disparities," Testimony for U.S. Commission on Civil Rights Briefing "The School to Prison Pipeline: The Intersection of Students of Color with Disabilities" (Dec. 8, 2017) (Att. 1), and handout for meeting with DOE staff (Mar. 22, 2019 (Att. 2) (pertaining to disparities in discipline and other educational outcome).

Many graphical and tabular illustrations of the pertinent patterns may be found in methods workshops given at University of Massachusetts Medical School (2015), UC Irvine (2015), George Mason University (2014), University of Maryland (2014), University Minnesota (2014), Harvard University (2012), and American University (2012). And many issues not addressed in any of the above works are addressed on the 100 plus pages and subpages of jpscanlan.com devoted to measurement/disparity issues.

I believe all of the above materials to be useful resources for government agencies attempting to assess equity issues. This response, however, merely suggests the scope of the matters that must be addressed for agencies to assess equity issues in a sound manner.

Section A outlines the key statistical issues with the focus described several paragraphs above. Section B discusses the pertinence of the issues addressed in Section A to cases involving barriers to voting with reference to the way those issues were highlighted in the recent Supreme Court decision in *Brnovich v. Democratic National Committee*. Section C discusses the impossibility of analyzing demographic differences based on a comparison of the proportion a group make up of the population with the proportion it makes up of persons experiencing an adverse or favorable outcome.

A. Patterns by Which Measures of Difference Involving Outcome Rates Tend to Affect by the Prevalence of an Outcome and the Mistaken Belief That Reducing a Barrier Will Tend Reduce, Rather Than Increase, Relative Demographic Differences in Failure to Overcome the Barrier.

Virtually all analyses of demographic differences involving favorable or corresponding adverse outcome have been undermined by a failure to understand patterns by which measures used to quantify demographic differences to be affected by the prevalence of an outcome. As a result of this failure, such analyses been unable to provide insight into whether the forces causing the outcome rates of advantaged and disadvantaged group to differ have increased or decreased over times or are larger in one setting than another or insight regarding the way policies affect such forces.

The pattern by which measures tend to be affected by the prevalence of an outcome that is most pertinent to subjects of the RFI is that whereby the rarer an outcome the greater tends to be the relative difference between rates at which advantaged and disadvantaged groups experience the outcome and the smaller tends to be the relative difference between rates at which the groups avoid the outcome (i.e., experience the opposite outcome). The matter could also be put in terms that whenever a favorable and corresponding adverse outcome change in prevalence, the relative difference for the increasing outcome tend to decrease while the relative difference for the decreasing outcome tends to increase. The matter can easily be illustrating with test score data showing that lowering a test cutoff – and thus making test passage more common and test failure less common – tends to reduce relative difference between the pass rates of higher- and lower-scoring groups while increasing relative differences between the groups' failure rates. The pattern can also be illustrated with myriad other types of data. It is also evident in what in fact commonly occurs when there is a change in the prevalence of a favorable and corresponding adverse outcome, especially when that change is substantial.

Nevertheless, even among persons who specialize in the analysis of demographic differences, virtually no one understand that it is even possible for the relative difference in a favorable outcome and the relative difference in the corresponding adverse outcome to change in opposite directions as the prevalence of an outcomes changes. And the overwhelming majority of such persons, and almost all federal agencies monitoring demographic differences for equity purposes, believe that generally reducing adverse criminal justice, school discipline, borrower, or health and healthcare outcomes will tend to reduce, rather than increase, relative differences in rates of experiencing the outcomes.

While there are many reasons why reducing an adverse outcome may not increase relative differences in the outcome in a particular situation, no one has ever advanced a reason why one should expect that reducing the prevalence of an adverse outcome would usually reduce relative differences in rates of experiencing the outcome. Rather, researchers and policymakers have merely taken the matter for granted. And among the countless observers who reinforce the mistaken belief by pointing out that a relative difference in an adverse outcome increased or persisted "despite" a general decline in the outcome, none appears to have considered that the repeated departures from an expectation might be evidence that the expectation is unsound.

Table 1 below shows the pass and fail rates of an advantaged group (AG) and a disadvantaged group (DG) at two cutoff points in a situation where the groups have normally distributed test scores with means that differ by half a standard deviation (a situation where approximately 31 percent of DG's scores are above the AG mean) and both distributions have the same standard deviation. The table also shows (in columns 5 through 7) three measures that might be used to quantify differences in test outcomes of AG and DG.

Table 1. Illustration of effects of lowering a test cutoff on measures of differences in test outcomes.

Row	(1) AG Pass Rate	(2) DG Pass Rate	(3) AG Fail Rate	(4) DG Fail Rate	(5) AG/DG Pass Ratio	(6) DG/AG Fail Ratio	(7) Abs Df (PP)
1	80%	63%	20%	37%	1.27	1.85	17
2	95%	87%	5%	13%	1.09	2.60	8

Column 5, which presents the ratio of AG's pass rate to DG's pass rate,[2] shows that at the higher cutoff, where pass rates are 80 percent for AG and 63 percent for DG, AG's pass rate is 1.27 times (27 percent greater than) DG's pass rate. If the cutoff is lowered to the point where AG's pass rate is 95 percent, DG's pass rate would be about 87 percent. At the lower cutoff, AG's pass rate is only 1.09 times (9 percent greater than) DG's pass rate. That lowering a cutoff tends to reduce relative differences in pass rates is well understood and underlies the widespread view that lowering a cutoff tends to reduce the disparate impact of tests on which some groups outperform others.

But, whereas lowering a cutoff tends to reduce relative differences in pass rates, it tends to increase relative differences in failure rates. As shown in column 6, initially DG's failure rate was 1.85 times (85 percent greater than) AG's failure rate. With the lower cutoff, DG's failure rate is 2.6 times (160 percent greater than) AG's failure rate.

Thus, lowering the cutoff, and making test passage more common and test failure less common, decreased the relative difference in the increasing outcome but increased the relative difference in the decreasing outcome. The same results would be observed if, instead of lowering the cutoff, education were improved sufficiently to all enable all persons falling between the two cutoffs to reach the higher the cutoff. This pattern holds across the entire range of possible test scores.

Column 7 shows that lowering the cutoff caused the absolute difference between the pass (and fail) rates of AG and DG to decrease from 17 to 8 percentage points. The decrease, however, is a function of the rate ranges at issue and would not be observed when, for example, a very high cutoff is lowered somewhat but remains quite high (and where the absolute difference would tend to increase), as reflected by movement from the first to the second rows of Table 5 of "Race and Mortality Revisited." While patterns of changes in absolute difference are not affected by whether one examines the favorable or the adverse outcome, it is affected by the prevalence of an outcome, though in a more complicated way than the two relative differences. Roughly, as uncommon outcome increase, absolute differences between rates of advantaged and disadvantaged group rates tend to increase, at least to the point where one group's rate reaches

---

[2] While I commonly refer to patterns of relative differences in this memorandum, the table actually presents ratios of two rates that I commonly refer to as rate ratio (and that are also termed risk ratios or relative risks). The relative difference is the rate ratio minus 1 where the rate ratio is above 1 and 1 minus the rate ratio where the rate ratio is below one. In the former case, the larger the rate ratio, the larger the relative difference; in the latter case, the smaller the rate ratio, the larger the relative difference. It may be more common to employ the disadvantaged group's rate as the numerator for ratios regarding the favorable as well as the adverse outcome, which is the approach as to favorable outcomes of the "four-fifths" or "80 percent" rule for identifying disparate impact under the Uniform Guideline for Employee Selection Procedures. I have sometimes employed this approach, as in "Can We Actually Measure Health Disparities?," *Chance* (Spring 2006). More recently, however, I have usually used the larger figure as the numerator for both ratios, in which case, as to both favorable and adverse outcomes, the larger the ratio, the larger the relative difference. Choice of numerator in the rate ratio, however, has no bearing on the patterns by which as the prevalence of an outcome changes, the two relative differences tend to change in opposite directions. I refer to patterns by which changes in the prevalence of an outcome affects relative differences rather than ratios because the accuracy of the statement turns on which figure is used as the numerator in the ratios.

50%. As the outcome further increases, the absolute difference tends to decrease, at least after the point where both groups rates have reached 50%.

The rates shown in the table are in the ranges usually found for many matters where demographic differences, including those involving school suspensions, arrests, loan rejections, foreclosures and evictions, unemployment, and poverty, are commonly quantified in terms of relative differences in the adverse outcomes. Increasingly, however, observers are also quantifying disparities for some of these matters in terms of absolute differences between rates, thus commonly reaching opposite conclusions from observers relying on relative differences. Few, however, appear to recognize that it is even possible for this to happen, much less that it will almost always happen when the referenced adverse outcomes change substantially.

But the absolute difference is not a main focus here. For the government does not mislead the public as to effects of reducing adverse outcomes on absolute differences between rates. It does so only with regard to relative differences in adverse outcomes. Thus, I will give only limited attention to the absolute difference here until discussing Table 5 infra.

Before turning to illustrations of pertinent patterns with actual data, I note that there are several things one should keep in mind with regard to the simple test score illustration in Table 1. First, federal agencies that promote beliefs that reducing an adverse outcomes will tend to reduce relative differences in rates of experiencing the outcomes have not reasoned as follows: while it is true that lowering a test cutoff and thus decreasing test failure will tend to increase relative differences in test failure rates, there are reasons why one should not expect a similar result from reducing other adverse outcomes and, in fact, should expect reducing other adverse outcome to reduce relative differences in rates of experiencing those outcomes. Rather, despite decades of dealing with racial differences in test outcomes, the agencies have yet to show an understanding even that lowering a test cutoff – or improving test performance or allowing the retaking of a test – will tend to increase relative differences between the failure rates of higher- and lower-scoring groups. Similarly, notwithstanding the connection between test scores and student proficiency, observers who analyze proficiency disparities in terms of relative difference in the favorable outcomes or the corresponding adverse outcomes have yet to realize that general improvements or general worsening of education, or changing to an easier or harder test, will tend to cause relative differences in meeting standards to change in opposite directions from relative differences in failure to meet the standard. See the CUNY ISLG Equality Indicators subpage of the Education Disparities page of jpscanlan.com.

Second, neither the increase in the relative difference in the adverse outcome, the decrease in the relative difference in the favorable outcome, nor the decrease in the absolute difference between rates effected by the lowering of the cutoff should be regarded as indicating that differences between the circumstances of AG and DG vis a vis the test have increased or decreased in any meaningful sense. Rather, none of the measures is a sound measure of association because each tends to be affected by the prevalence of an outcome. And if, for example, the two rows reflected the results of subjective judgments of decision-makers as to passing or failing a test or any other favorable and corresponding adverse outcomes, there would be no basis for maintaining that either of the decision-makers was more likely to have engaged in discrimination than the other.

It is useful to keep in mind, however, that if the pass and fail rates are the results of a subjective judgments of different teachers, other things being equal, teachers who are more lenient graders or more effective teachers than other teachers will tend to show patterns more like that in Row 2 than Row 1, while other teachers will tend to show patterns more like that in Row 1 than Row 2. Similarly, if the adverse outcome rates in the table are rates of suspension from school for particular teachers, teachers who are more lenient disciplinarians, better able to maintain discipline without resort to suspensions, or more responsive to encouragements to generally reduce suspensions than other teachers will tend to show patterns more like that in Row 2 than in Row 1, while other teachers will tend to show patterns more like that in Row 1 than Row 2. And when the adverse outcome rates involve the use of force in making arrests, officers or agencies that are more circumspect about the use of force and better skilled at de-escalation techniques will tend to show patterns more like that in Row 2 than Row 1 than other officers and agencies, while other officers and agencies will tend to show patterns more like that in Row 1 than Row 2. Thus, other things being equal, decision-makers whose conduct most accords with that which governments encourage may face the greatest chances of being accused of discrimination on the basis of the comparative size of relative racial difference in adverse outcomes resulting from their actions.

Third, when presented with the actual underlying rates, astute observers may recognize the extent to which observed patterns of measures of differences between rates may be functions of the prevalence of an outcome. They may also recognize that different measures are yielding opposite conclusions as to the comparative size of a demographic difference over time or in one setting than another, a clue as to potential unsoundness of one or all of the measures. But, too often, regardless of the measure being employed to quantify a demographic difference, the measure is presented without the underlying rates. Thus, it is crucial that presentations on demographic difference include actual rates at which groups experience an outcome.

Finally, any reduction in the forces causing the outcome rates of advantaged and disadvantaged groups to differ – including racial or other bias when that is a cause of such differences – will cause all measures of differences between outcome rates to be smaller than they would otherwise be. But it will be impossible to determine whether those force have increased or decreased over time, or the efficacy of policies aimed at reducing those forces, without understanding the ways measures of difference between outcomes rates are also being affected by changes in the prevalence of an outcome. For example, even though a program has reduced or eliminated the role of racial bias with respect to some outcome, general reductions in the prevalence of an outcome may cause the relative racial difference to increase.

Table 2 is based on data from a Department of Education (DOE) study that showed rates at which demographic groups fell into various levels of prose, document, and quantitative literacy. Following the format of Table 1, the table shows the white and Black rates of reaching (Fav Rate) and failing to reach (Adv Rate) certain levels of each type of literacy, along with the same measures of differences for favorable and adverse outcome rates shown in Table 1. And the table shows that the lower the level on which literacy might be appraised, the smaller is the relative difference in reaching the level while the larger is the relative difference in failure to reach the level. Moving down the rows in each section, the final column shows that the absolute difference increase as the favorable outcome (reaching a particular level) goes from being uncommon to being somewhat common and decrease as the favorable outcome increases from being somewhat common to being very common.

Table 2. White and Black Rates of reaching and failing to reach various levels of prose, document, and quantitative literacy, with measures of difference.

Prof Type	Level	(1) White Fav Rt	(2) Black Fav Rt	(3) White Adv Rate	(4) Black Adv Rate	(5) Wh/Bl Fav Ratio	(6) Bl/Wh Adv Ratio	(7) Abs Df (PP)
Prose	Proficient	17%	2%	83%	98%	8.50	1.18	15
Prose	Intermediate	68%	23%	32%	77%	2.96	2.41	45
Prose	Basic	93%	76%	7%	24%	1.22	3.43	17
Document	Proficient	15%	2%	85%	98%	7.50	1.15	13
Document	Intermediate	73%	42%	27%	58%	1.74	2.15	31
Document	Basic	92%	76%	8%	24%	1.21	3.00	16
Quantitative	Proficient	17%	2%	83%	98%	8.50	1.18	15
Quantitative	Intermediate	44%	17%	56%	83%	2.59	1.48	27
Quantitative	Basic	87%	53%	13%	47%	1.64	3.62	34

The illustration is particularly pertinent to the barrier issues that are the principal concerns of the RFI. For it shows how an agency's making its communications regarding its programs easier to understand will tend to reduce relative differences between rates at which advantaged and disadvantaged groups understand such communications and utilizes the programs but increase relative differences between rates at which such group fail to understand such communications and fail to utilize the programs. Similarly, the more agencies simplify the process of participating in its programs, the smaller will tend to be relative differences between rates at which advantaged and disadvantaged groups participate in the programs, but the larger will tend to be relative differences in rates at which the groups fail to participate in the programs.

Tables 3 and 4 present Black and white rates of falling above and below various income or credit score levels, with the same measures of relative differences in favorable and adverse outcomes shown in the prior tables.[3]

Table 3. Illustration of effects of lowering an income requirement on relative. differences in meeting the requirement and relative differences in failing to meet the requirement.

Income	(1) Perc of Wh Abv	(2) Perc of Bl Abv	(3) Perc of Wh Bel	(4) Perc of Bl Bel	(5) Wh/Bl Abv Ratio	(6) Bl/Wh Bel Ratio
$100,000	27.0%	12.1%	73.0%	87.9%	2.23	1.20
$85,000	34.6%	17.3%	65.4%	82.7%	2.00	1.26
$75,000	41.1%	22.7%	58.9%	77.3%	1.81	1.31
$60,000	52.5%	31.3%	47.5%	68.7%	1.68	1.45
$50,000	61.0%	39.2%	39.0%	60.8%	1.56	1.56

[3] The sources of the data are explained on the Income and Credit Score Examples subpage of the Lending Disparities page of jpscanlan.com.

Table 4. Illustration of effects of lowering a credit score requirement on relative differences in meeting the requirement and relative differences in failing to meet the requirement.

Score	(1) Perc of Wh Abv	(2) Perc of Bl Abv	(3) Perc of Wh Bel	(4) Perc of Bl Bel	(5) W/B Abv Ratio	(6) B/W Bel Ratio
740	46.80%	19.50%	53.20%	80.50%	2.40	1.51
720	57.77%	27.01%	42.23%	72.99%	2.14	1.73
700	67.83%	35.67%	32.17%	64.33%	1.90	2.00
680	76.73%	45.42%	23.27%	54.58%	1.69	2.35
660	83.90%	55.70%	16.10%	44.30%	1.51	2.75

The tables show how lowering an income or credit score requirement, while tending to reduce relative racial differences in rates of meeting the requirement, tends to increase relative racial differences in rates of meeting the requirement. Despite the availability of such data, all federal agencies enforcing federal fair lending laws have long operated on the belief that relaxing lending standards will tend to reduce relative racial differences in home mortgage and other loan rejection rates. See, e.g., "Race and Mortality Revisited," "The Perverse Enforcement of Fair Lending Laws," *Mortgage Banking* (May 2014); "Misunderstanding of Statistics Leads to Misguided Law Enforcement Policies," *Amstat News* (Dec. 2012). The same issues apply to the government's administration of its own programs.

Income data of the type presented in Table 3 would also show how general reductions in poverty tend to increase relative racial differences in poverty rate at the same time the reductions reduce relative racial differences in rates of avoiding poverty. It should be evident that efforts to study how programs or general trends affect racial differences regarding poverty can provide nothing of value, though much that is misleading, if conducted without consideration of the way the measures to quantify the differences are being affected by general changes in poverty rates. See discussion of Table 2 (at 329-330, 343) of "Race and Mortality Revisited."

Further regarding the housing issues as to which the data in Tables 3 and 4 are especially pertinent, disadvantaged groups are more likely to have difficulty meeting loan and rental payments than advantaged groups. The matter is commonly cast in terms of the ratio of distress rates of a disadvantaged groups to those of advantaged groups.[4] But differences would be evident regardless of the measure used to quantify them. There exists a universal belief that government actions that reduce foreclosures and evictions will tend to reduce relative racial and other differences in foreclosures and evictions. In fact, however, such measures tend to increase

[4] The greater difficulty for disadvantaged groups is the case regardless of the COVID-19 pandemic. Many have attributed large relative racial/ethnic differences in evictions/foreclosures to the COVID-19 pandemic. But relative differences in these outcomes were probably larger, while relative in rates of avoiding these outcomes were probably smaller, before the pandemic led to general increases in those outcomes.

relative differences in foreclosures and evictions, while reducing relative differences in rates of avoiding these outcomes.[5]

Table 5 (a version or Table 4 in "Race and Mortality Revisited") is based on a 2008 *Pediatrics* study that examined the effect of a school-entry hepatitis B vaccination requirement on racial and ethnic differences in vaccination rates. The table employs the same formatting as Table 1, but with an additional column that I will explain below.

Table 5. White and black Hepatitis B vaccination rates in grades 5 and 9 before and after imposition of school-entry vaccination requirement, with measures of difference

Grade	Year	Req	Wh Vac Rt	Bl Vac Rt	Wh No Vac Rt	Bl No Vac Rt	W/B Ratio Vac	B/W Ratio No Vac	Abs Df	EES
5	1996	Pre	8%	3%	92%	97%	2.67	1.05	5PP	0.47
5	1997	Post	46%	33%	54%	67%	1.39	1.24	13PP	0.34
9	1996	Pre	46%	32%	54%	68%	1.44	1.26	14PP	0.37
9	1997	Post	89%	84%	11%	16%	1.06	1.45	5PP	0.24

The table shows the usual pattern whereby general increases in the favorable and general decreases in the corresponding adverse outcome are accompanied by reduced relative differences in the favorable and increased relative differences in the adverse outcome. The penultimate column also shows the usual pattern whereby when an uncommon outcome became somewhat common (as in grade 5) the absolute difference increases and when a somewhat common outcome became quite common (as in grade 9) the absolute difference decreases.

The authors measured racial disparities in terms of relative differences in vaccination rates and found that the dramatic increases in vaccination rates had been accompanied by substantial decreases in disparities. Many disparities researchers (including those at the Centers for Disease Control and Prevention) measure vaccination disparities in terms of absolute differences between rates, while rarely if ever suggesting that others might measures them differently. The researchers who measure disparities in terms of absolute differences between rates would say that the disparities had increased in grade 5 but decreased in grade 9.

Table 5 also shows the types of patterns lately observed as COVID-19 vaccination rapidly increased from being very uncommon to being somewhat common and now is further increasing to being very common. That is, relative difference in receipt of vaccination will tend

[5] See The Bureau of Financial Protection June 30, 2021 Final Rule "Protections for Borrowers Affected by the COVID-19 Emergency Under the Real Estate Settlement Procedures Act." After describing relative racial/ethnic differences in rates of being behind on mortgage payments or in forbearance (cast in terms of the ratio of Black and Hispanic rates to the white rates), the rule states (FR 34889): "The benefit to avoiding foreclosure for these arguably 'marginal' borrowers may be significantly larger compared to the average borrower." This statement is correct in that the general reductions in foreclosures presumably effected by the rule will tend to increase the Black and Hispanic rates of avoiding foreclosure proportionately more than the average (and white) borrower. But it will tend to reduce foreclosure rates proportionately more for the average (and white) borrower than Black and Hispanic borrowers. That would increase relative racial/ethnic differences in foreclosure rates at the same time that it reduces such differences in rate of avoiding foreclosure. Sound appraisals of the value of the rule with regard to racial/ethnic disparities must be informed by an understanding of such patterns.

to decrease, and relative differences in nonreceipt vaccination will tend to increase, throughout the entire process. Absolute difference between rates, however, will tend to increase for a time and then decrease. Effort to determine the utility of programs aimed at reducing vaccination hesitancy among particular groups must be informed an understanding of such patterns.

I typically use the data in Table 5 to illustrate the general disarray in health and healthcare disparities research, especially within the government. As discussed in "Race and Mortality Revisited" (at 332-335) and throughout "The Mismeasure of Health Disparities," as early as 2004 National Center for Health Statistics (NCHS) statisticians recognized that improvements in health and healthcare tend to reduce relative differences in the increasing (favorable) outcomes like survival and receipt of appropriate care, while increasing relative differences in the corresponding (decreasing) adverse outcomes like mortality and nonreceipt of appropriate care. Given that the forces that cause favorable outcome rates of advantaged and disadvantaged groups to differ are exactly the same forces that cause the corresponding adverse outcome rates to differ, one would expect that NCHS's recognition of the described pattern would cause the agency to question whether either of the two relative differences was a useful indicator of whether such forces are increasing or decreasing over time.

Instead, however, the agency decided that, for purposes of appraising progress in reaching of Healthy People 2010 disparities reduction goals, it would measure all disparities in terms of relative differences in adverse outcomes (which, in the case of healthcare, would be nonreceipt of appropriate care). But the agency also decided that where indicators had previously been cast in favorable terms, like receipt of vaccination, it would continue to cast the indicators in the favorable terms even though disparities were being measured in terms of relative differences in the corresponding adverse outcomes.

Thus, in the case of the pattern in Table 5, on the basis of the fact that relative racial differences in nonreceipt of vaccination increased in grades 5 and 6 following the requirement, NCHS would have found that racial disparities in vaccination rates had increased. As discussed in "The Mismeasure of Health Disparities" (at 418), however, in 2015 the agency changed that recommendation. For purposes of appraising progress in reaching Healthy People 2020 disparities reduction goals, it would measure disparities in things like vaccination in terms of relative difference in receipt of vaccination.

The Agency for Healthcare Research and Quality (AHRQ), which produces the yearly National Healthcare Disparities Reports National Healthcare Disparities Reports (NHDR) (which since 2013 have been part of the National Healthcare Quality and Disparities Reports), followed NCHS guidance both with respect to measuring healthcare disparities in terms of relative differences in nonreceipt of care and with respect to describing indicators in favorable terms even though the disparity was being measuring in terms of relative difference in nonreceipt of care. AHRQ has never changed that approach. And it has never shown a recognition that it even possible for the relative differences in a favorable outcome and the relative difference in the corresponding adverse outcome to change in opposite directions as the prevalence of an outcome changes.

AHRQ did, however, change its approach to measurement of health and healthcare disparities in an important respect beginning with the 2010 NHDR. Prior to that report, while measuring disparities in terms of relative differences in adverse outcomes, the agency quantified changes in disparities in terms of percentage point changes in relative differences. That is, for example, with respect to the data on grade 5 in Table 5, AHRQ would regard the disparity to have changed from a 5% relative difference to a 24 % relative difference in nonreceipt of vaccination and would consider that a 19 percentage point increase in the relative difference.

From the 2010 report on, however, while continuing to quantify disparities in terms of relative differences in adverse outcomes, AHRQ made determination about the directions of changes in disparities on the basis of the comparative size of absolute changes in the rates of the groups being compared (which is the same thing as comparing the size of absolute differences at the beginning and end of the period analyzed). Under this approach, the agency would still regard the disparity to be larger after than before imposition of the requirement, both for grade 5 and grade 9. It would also regard the disparity to have increased in grade 5 because the absolute difference increased from 5 to 13 percentage points. But in the case of grade 9, even though it would regard the disparity to be larger after than before imposition of the requirement, the agency would regard disparity to have decreased over time.

In the 2012 report, this approach caused the agency to highlight as some of fastest reductions in healthcare disparities over a particular period (on the basis of the comparative size of absolute change in the rates of the groups being compared) situations where the agency also would regard the disparities to be much larger at the end of the period than at the beginning of the period (on the basis of the comparative size of relative differences in nonreceipt of care). The situations of this nature that caught my attention all involved instances of high vaccination rates that grew even higher during the period examined. They may be found in Table 6 (at 21) of my July 1, 2015 letter explaining that and other problems with the NHQDRs to the agency.

As discussed in the NHDR Measurement Issues subpage of the Measuring Health Disparities page of jpscanlan.com, AHRQ attempted to understand this issue after I brought it to the agency's attention. But whether it was able to do so or not, the agency has continued this approach. We may well see manifestations of the approach in the agency's analysis of demographic differences regarding COVID-19 vaccination or any other COVID-19-related favorable and corresponding adverse health and healthcare outcomes.

Various misunderstandings about quantification of demographic differences specifically regarding COVID-19-related matters are discussed at several places in the HUD Panel Mem. (Att. 3), especially pages 9-10. These include the failure to understand why relative racial differences in adverse COVID-19 outcomes are commonly (or always) greater, while relative differences in the corresponding favorable outcomes are commonly (or always) smaller, among younger persons and other advantaged subpopulations than the corresponding disadvantaged subpopulation, as well as the failure to understand that improvements in care of COVID-19 patients will tend to increase relative racial differences in morality among such patients at the same time that improvement reduce relative racial differences in survival among such persons.

The situation at AHRQ is a striking example of problems in health and healthcare disparities research. But the typical such research that reflects no understanding of how the measure it employs tends to be affected by the prevalence of an outcome or that different measures tends to be affected differently by the prevalence of an outcome is equally capable of misleading observers. Without understanding such things, researcher cannot sensibly discuss the effects of policies or anything else on differences in the health and healthcare-related circumstances of advantaged and disadvantaged groups.

The final column of Table 5 shows the measure described in "Race and Mortality Revisited" and most of the longer items mentioned in the Introduction (and employed in each of the attachments) that is theoretically unaffected by the prevalence of an outcome. I commonly term the measure EES for estimated effect size and statisticians may refer to it as probit d. It involves deriving from a pair of favorable or corresponding adverse outcomes the difference between means of underlying normal risk distributions. The figures in the column are the differences between those means in terms of percentage of a standard deviation.

The modest decreases in the measure for both grades accords what seems reasonable to expect in the case of a mandated requirement. For such requirements would seem to counter somewhat the strength of the forces causing outcome rates of advantaged and disadvantaged groups to differ. But the effect was insufficient to cause departures from the usual patterns of changes in standard measures.

The described approach to measurement is imperfect, as I discuss, for example, in "Race and Mortality Revisited" (at 337). But it is vastly superior to relying on any measure without consideration of way the measures employed tend to be affected by the prevalence of an outcome.[6] This response, however, is largely limited to explaining problems with usual methods of analyzing equity issues without consideration of the effects of the prevalence of an outcome on those methods, and sometimes while mistakenly believing that reducing the prevalence of an outcome will tend to reduce a measure of equity when in fact it will tend to increase the measure.

B. The Courts' Failure to Understand the Ways Measures of Demographic Differences Tend to Be Affected by the Prevalence of an Outcome.

Courts have commonly analyzed discrimination issues, whether characterized in terms of disparate impact of disparate treatment, in terms of relative differences in favorable outcomes or relative differences in adverse outcomes. To my knowledge, however, no court has recognized that it is even possible for the two relative differences to change in opposite directions, even when a court has discussed how the size of a relative difference turns on whether one examines the favorable or the corresponding adverse outcomes. See Jones v. City of Boston subpage of the Disparate Impact page (DIP) of jpscanlan.com. Presumably, all federal and state judges who have any view on the matter share the mistaken view that relaxing a standard tends to reduce

[6] Probit d' was recently treated as a robust measure of school suspension disparities in Erik J. Girvan, Kent McIntosh & Keith Smolkowski, "Tail, Tusk, and Trunk: What Different Metrics Reveal About Racial Disproportionality in School Discipline," Educational Psychologist (2019).

relative differences in failure to meet it. The Fisher v. Transco Services subpage of DIP discusses a 1992 case reflecting that view. I do not believe anything has changed since 1992.

The recent case of *Brnovich v. Democratic National Committee*, No. 19-1257 (July 1, 2021), where by a 6-3 margin the Supreme Court rejected a challenge to certain Arizona voting regulations, may usefully illustrate the current understanding of the issue among the Justices. All Justices, like all counsel and expert witnesses for the litigants, presumably share the near universal view that relaxing a requirement would tend to reduce, rather than increase, relative racial/ethnic difference in failure to meet the requirement.[7]

With respect to the quantification of the racial/ethnic impact of a requirement, a crucial issue in the majority's analysis, the majority declined to rely on relative differences in failure to meet the requirement for purposes of that quantification. It did so while regarding reliance on a very large relative difference for a rare adverse outcome a type of manipulation of statistics. Op. 28. But it did so without evident awareness that, regardless of the frequency of an outcome, neither the relative difference in the adverse outcome nor the relative difference in the corresponding favorable outcome is a sound measure of association because each tends to be systematically affected by the prevalence of an outcome. The majority opinion also showed no awareness that generally reducing an outcome would in fact tend to increase relative differences in rates of experiencing the outcome. The majority may have been even more inclined to reach the result it did if it were aware that making a requirement less stringent would tend to increase relative racial/ethnic differences in rates of failure to meet the requirement.

The dissent, which strongly favored the use of relative difference in failure to meet a requirement to quantify the impact of the requirement, seemed plainly of the view that the more stringent requirements tended to result in larger, rather than smaller, relative difference in failure to meet the requirements. The dissent also found historical significance in the fact that across the South the Australian Ballot had caused a larger proportionate reduction in voter participation rates among African Americans than among whites. Dis. Op. 28. But the above-described pattern by which relative differences in favorable and corresponding adverse outcomes tend to change as the prevalence of an outcome is simply a reflection of the fact that anytime there occurs a general change in the prevalence of an outcome for two groups with different baseline rates for the outcome, the group with the lower baseline rate for the outcome will tend to have the larger proportionate change in its rate for the outcome than the other group, while the other

[7] I have previously discussed the failure to understand that relaxing a requirement tends to increase relative difference in failure to meet the requirement specifically with reference to voting rights case in "Misunderstanding of Statistics Confounds Analyses of Criminal Justice Issues in Baltimore and Voter ID Issues in Texas and North Carolina," Federalist Society Blog (Oct. 3, 2016), "Will Trump Have the First Numerate Administration?" Federalist Society Blog (Jan. 4, 2017). I do not know the extent to which all or most voter rights cases involve the issues discussed here and know little about the recently-filed case of *US v. The State of Georgia*, No. 1:21-cv-02575 (N.D. Ga.). But the complaint in the case states (at 8): "According to the 2015-2019 American Community Survey 5-Year Estimates, Black households in Georgia were more than three times as likely as non-Hispanic white households to lack access to a vehicle (12.9% with access compared to 3.9%)." To the extent that transportation to voting sites is an issue, that would seem another matter where reducing the burden of finding transportation to a voting site, while tending to reduce relative racial differences in getting to voting sites, would tend to increase relative racial differences in failure to get to sites – just as making vehicles easier to secure in the state would tend to reduce relative racial differences in access to a vehicle, while increase relative differences in lacking access to a vehicle.

group (which is the group with the lower baseline rate for the opposite outcome) will tend to have the larger proportionate change in its rate for the opposite outcome than the first group. See "Race and Mortality Revisited" (at 339). See also note 5 supra. Raising the cutoff to the original position in Table 1, for example, would cause DG to experience a larger proportionate decrease in its pass rate than AG (27.6% for DG versus 15.8% for AG), while causing a larger proportionate increase in failure rates for AG than DG (300% for AG versus 184.6% for DG).

Assuming African Americans had lower baseline voter participation rates than whites before introduction of the Australian ballot, a larger proportionate reduction for African Americans than for whites is to be expected for that reason alone, as would be a larger proportionate increase in non-participation rates among whites than African Americans. But one would need to know that actual rates (and know them for each state) in order to know whether the Australian ballot affected one group more than another in any meaningful sense.

As long as the impact of policies is appraised in terms of relative difference in favorable or adverse outcomes (something the *Brnovich* majority opinion may have left open in cases other than where one outcome is rare and the other nearly universal), courts must eventually consider, with respect to requirement that entities implement less discriminatory alternative for even policies that can be justified, whether relaxing a standard decreases the adverse impact (measured by relative difference in meeting the standard) or increases the impact (measured by relative differences in failure to meet the standard). When courts do consider the issue, they will search in vain for an indication of Congressional intent on such matters, since no legislators involved in imposing the less discriminatory alternative requirement (as in the Civil Rights Act of 1991) were aware that it was even possible for lowering a test cutoff to increase relative differences in failure rates at the same time that it decreases relative differences in pass rates. See generally my "Is the Disparate Impact Doctrine Unconstitutionally Vague?" Federalist Society Blog (May 6, 2016), Kansas Law Paper (at 37-32), TDHCD Brief (at 20-23).

C. The Impossibility of Analyzing Demographic Differences by Comparing the Proportion a Group Makes Up of a Population with the Proportion It Makes Up of Persons Experiencing a Favorable or Adverse Outcome.

This section discusses the impossibility of quantifying a demographic difference on the basis of a comparison between the proportion a group makes up of a population and the proportion it makes up of persons experiencing an outcome. An overriding point is that any presentation of data on group differences must present the actual rates at which the groups experience the outcomes. Another crucial point, however, is that whereas government agencies and many others promote the view that generally reducing certain adverse outcomes, including criminal justice and school discipline outcomes, will tend to reduce the proportions Blacks make up of persons experiencing the outcomes, in fact reducing such outcomes tends to increase such proportions in the same way that it tends to increase relative racial differences in rates of experiencing the outcomes. This was an important point of my testimony before the Commission on Civil Rights (Att. 1).

Many equity issues are analyzed by comparing the proportion a subject group makes up of a population (PP) with the proportion it makes up of person experiencing a favorable outcome

(PFO) or adverse outcome (PFO). This is often the approach for adverse criminal justice and school discipline outcomes, as well as many favorable educational outcomes, and has lately even been the approach with respect to COVID-19-related favorable or adverse health and healthcare outcomes. This is an unsound method of assessing inequity for the same reasons that the two relative differences and the absolute difference between rates are unsound measures of demographic differences. Specifically, PFO and PAO tend to be systematically affected by the prevalence of an outcome and hence any measure of difference between PP and PFO or PAO tends to be systematically affected by the prevalence of the outcome.

Table 6 below is similar to Table 1, except that absolute difference between rates is omitted and the final two columns present the proportion DG makes up of persons who pass the test and persons who fail the test at each cutoff, in circumstances where DG comprises 50% of test takers.

Table 6. Illustration of effects of lowering a test cutoff on measures of differences in test outcomes of advantaged group (AG) and disadvantaged group (DG) (where DG comprises 50% of test takers).

Row	(1) AG Pass Rate	(2) DG Pass Rate	(3) AG Fail Rate	(4) DG Fail Rate	(5) AG/DG Pass Ratio	(6) DG/AG Fail Ratio	(7) DG Prop of Pass	(8) DG Prop of Fail
1	80%	63%	20%	37%	1.27	1.85	44%	65%
2	95%	87%	5%	13%	1.09	2.60	48%	72%

The final columns show that lowering the cutoff, and thus making test passage more common and test failure less common, increases the proportion DG makes up both of persons who pass the test (from 44% to 48%) and persons who fail the test (from 65% to 72%). Because the proportion DG makes up of persons taking the test remains unchanged, lowering the cutoff would reduce all measures of difference between the proportion DG makes up of persons who take the test and persons who pass the test and increase all measures of difference between the proportion DG makes up of persons who take the test and persons who fail the test. While the hypothetical posits that DG is 50% of test takers in order to illustrate the pattern, the pattern of directions of changes in the proportions DG makes up of persons who pass the test and persons who fail the test would hold regardless of the proportion DG makes up of persons who take the test.

When there are only two groups in the population being examined, directions of change in the size of measures of difference between PP and PFO and between PP and PAO are the same as the directions of changes in relative differences in the associated favorable and adverse outcomes (though the matter becomes more complicated when the population is comprised of more than two groups[8]). For that reason alone, any difference between PP and either PFO or PAO would be an unsound measure of association for the same reasons that the two relative difference between rates are unsound measures of association.

[8] For this reason, I sometimes discuss the matter in terms of the way changing the prevalence of an outcome affected the proportion Blacks make up of the combined Black and white populations, as in Table 1 of "Can We Actually Measure Health Disparities?," *Chance* (Spring 2006).

Further, information on PP and PAO or PFO allows one to determine the relative differences between the rate at which the subject group experiences whichever of the two outcomes for which is information is presented and the rate at which all other persons experience the outcome. But it does not enable one to determine the actual rates at which each group experiences the outcomes (and hence the relative differences in the corresponding adverse outcome or the actual difference between rates, and, more important, a measure like EES).[9] One must know the actual rates at which groups being compared experience an outcome in order to quantify the strength of the forces causing the favorable and corresponding adverse outcomes of the two group to differ or even to know the extent to which the proportion the group makes up of persons experiencing the outcome may be a function of the prevalence of an outcome.

But even persons who believe that either of the two relative differences (or the absolute difference) can effectively quantify the difference between the circumstances of the two group vis a vis the outcome and its opposite should recognize other reasons why comparing PP and PFO or PAO is an absurd way to quantify demographic differences. For one thing, when there are more than two groups in the population being examined, some groups may be found to be underrepresented among persons experiencing an adverse outcome or overrepresented among persons experiencing the favorable outcome, even though the groups' favorable and adverse outcome rates are worse than the rates for an advantaged group. For example, Hispanic students, whose rates for suspensions from school are usually greater than the rates for white students but lower than rates for Black students, are commonly regarded as underrepresented among suspended students even though their rates are higher than the rates of white students. And even when general reductions in suspensions are accompanied by increases in relative differences between Hispanic and white suspension rates, in places where Black students make up a significant proportion of students, the proportion Hispanic students make up of suspended students may decrease (though the proportion Hispanic students make up of Hispanic and white students combined has increased).

Even when there are two only racial/ethnic groups in the population being examined, however, there are evident absurdities in quantifying demographic differences on the basis of comparison between PP and either PAO or PFO. One aspect of the problem is also pertinent to any comparisons between a group's outcome rate and the overall rate and may be more easily explained in that context. For the subject group's own rate influences the overall rate with which the group's rate is being compared. And the greater the proportion a group makes up of the population, the greater will be the influence of its rate on the overall rate, thus reducing all measures of difference between the group's rate and the overall rate. That is why it never makes sense to compare a disadvantaged group's rate with an overall rate rather than the rate of an advantaged group even when there are only two groups in the population.

---

[9] Knowing only that DG is 50% of the population (DGPP) and 65% of persons experiencing the adverse outcome of test failure (DGPAO) (as in Row 1 of Table 6),one can determine from the formula ([DGPAO]/[DGPP])/((1-[DGPAO])/(1-[DGPP])) that the ratio of DG's failure rate to AG's failure rate is 1.85, the same ratio that would be calculated from the actual failure rates if they are known. But one cannot make sense out of the 1.85 ratio in the way a numerate observers might make sense out of the underlying failure rates (which also reveal the underlying pass rates).

Comparisons of PP with PAO or PFO are necessarily comparisons of the group's situation with an overall situation that is influenced by the group's own situation. But there are additional problems with comparing PP and PAO or PFO arising from the ways in which differences between and PAO or PFO are quantified. The May 2014 document "Methods for Assessing Racial/Ethnic Disproportionality in Special Education," which was funded by the DOE and produced by an arm of IDEA Data Center (a part of Westat, Inc.) recommends measuring differences between PP and PAO for assignment to special education or discipline in either relative or absolute terms. The April 2018 Government Accountability Office report titled "Discipline Disparities for Black Students, Boys, and Students With Disabilities" measured discipline disparities in terms of the absolute differences between the proportion Black students (or male students or students with disabilities) made up of students and the proportion they made up of students experiencing an adverse discipline outcome (PAO – PP), as did the DOE for certain matters in its April 2018 document "Data Highlights on School Climate and Safety in Our Nation's Public Schools."

In the DOE's June 2021 document "An Overview of Exclusionary Discipline Practices in Public Schools for the 2017-2018 School Year" and in its June 8, 2021 Request for Information Regarding the Nondiscriminatory Administration of School Discipline, the agency measured discipline disparities in terms of relative differences between the proportion a group or subgroup makes up of students and the proportion it makes up of students experiencing an adverse discipline outcome – that is, (PAO – PP)/PP or (PAO/PP) -1), while sometimes characterizing the matter in terms of the ratio of PAO to PP.

At the same that an increase the proportion a group makes up of persons experiencing the outcome affects the situation of the overall population with whose situation the group's situation is being compared, the increase also affects relative differences and absolute differences between PP and PAO, and it does so in conflicting ways, as illustrated in Table 7.

Table 7. Relative and absolute differences between proportion Black students make up of all students and proportion Black students make up of suspended students in schools where Black and white suspension rate are 15% and 5% with various Black proportions of all students.

Black Rate	Black Prop Students	Black Prop Suspensions	Abs Df (PP)	Rel Df
15%	5%	13.64%	8.64	172.73%
15%	20%	42.86%	22.86	114.29%
15%	40%	66.67%	26.67	66.67%
15%	60%	81.82%	21.82	36.36%
15%	80%	92.31%	12.31	15.38%
15%	95%	98.28%	3.28	3.45%

Table 7 presents situations where all students are either Black or white and Black and white suspension rates are 15% and 5%, though the proportion Black students make up of students varies. There is no basis for distinguishing between schools that have those rates with the respect to the magnitude of the difference (though one might find varying reasons why Black and white rates differ in each situation). Consider what the table illustrates regarding the way the proportion Black students make up of all students affects absolute and relative differences between PP and PAO when the Black proportion of students increases incrementally from 5% to 95%.

The relative difference between PP and PAO decreases consistently from the point where Black students make up 5% of students (and PAO is 172.7% greater than PP) to the point where Black students make up 95% of students (and PAO is only 3.24% greater than PP). The absolute difference between PP and PAO, however, is affected by increases in PP in a more complicated way. It is 8.64 percentage points when Black students make up only 5% of students, increases until reaching a maximum of 26.7 percentage points when Black students make up approximately 40% of students, and then declines until reaching 3.28 percentage points when Black students make up 95% of students.

Not only do both approaches find differences in disparities from school-to-school even though the situation is exactly the same in each school, the two approaches can yield different view as to which of two schools has the larger disparity problem. Notice that in the schools where Black students make up 20% and 40% of students, the former school shows the larger relative difference while the latter school shows that larger absolute difference.

The same issues exist when racial difference in favorable outcomes are examined by comparing the proportion Black students make up of students (PP) with the proportion they make up of students experiencing a favorable outcome (PFO) like assignment to a gifted and talented program. Table 8 presents a situation where gifted and talented rates are 5% for Black students and 15% for white students. The negative signs in the final two columns reflect the fact that the values show the degree to which the Black PFO is less than the proportion Black students make up of all students.

Table 8. Relative and absolute differences between proportion Black students make up of students and proportion Black students make up of students in gifted and talented programs in schools where Black and white rates of assignment to programs are 5% and 15% with various Black proportions of all students.

Black Rate	White Rate	Bl Prop Students	Bl Prop GIFT	Ab Df (PP)	Rel Df
5%	15%	5%	1.72%	-3.28	-65.52%
5%	15%	20%	7.69%	-12.31	-61.54%
5%	15%	40%	18.18%	-21.82	-54.55%
5%	15%	60%	33.33%	-26.67	-44.44%
5%	15%	80%	57.14%	-22.86	-28.57%
5%	15%	95%	86.36%	-8.64	-9.09%

The table shows patterns similar to those in Table 7, though the absolute and relative difference values are somewhat different from those in Table 7 and Black students are a larger proportion of students (approximately 63%) at the point where the percentage point difference reaches a maximum (in negative terms). As in the prior example respecting schools with 20% and 40% of students, the former school shows the larger relative difference while the latter school shows the larger absolute difference.

Now consider how this approach might cause observers to find suspension differences to be larger in one school than another when they are in fact smaller. School A has the same Black and white 15% and 5% rates suspension rate as those in the earlier example, while School B has Black and white suspension rates of 13% and 5%. Because the white rate is the same in the two situations all measures of difference between Black and white rates would be larger in School A than in School B.

In this situation, as shown in Table 9, even though the racial difference is greater in School A than School B, if Black students were 10% of students in School A and 20% of students in School B, the absolute difference between PP and PAO would be greater in School B than School A (19.39 versus 15.0 percentage points). On the other hand, if Black students were 20% of students in School A and 10% of students in School B, the relative difference between PP and PAO would be greater in School B than School A (124.14% versus 114.97%).

Table 9. Relative and absolute differences between proportion Black students make up of all students and proportion Black students make up of suspended students in schools where Black and white suspension rate are 15% and 5% (School A) or 13% and 5% (School B) with various Black proportions of all students of 10% and 20%.

School	Black Rate	White Rate	Black Prop Students	Black Prop Suspensions	Abs Df (PP)	Rel Df
A	15%	5%	10%	25.00%	15.00	150.00%
A	15%	5%	20%	42.86%	22.86	114.29%
B	13%	5%	10%	22.41%	12.41	124.14%
B	13%	5%	20%	39.39%	19.39	96.97%

As noted at the outset of this section, both adverse and favorable COVID-19-related health and healthcare outcomes have lately been analyzed in terms of comparisons of PP and PAO or PFO for various demographic groups. Sometimes the differences have been quantified in terms of relative differences between PP and PAO or PFO and sometimes they have been quantified in terms of absolute differences between PP and PAO or PFO. And many comparisons are made across states that have different racial/ethnic makeups. Thus, these efforts to compare the size of demographic differences implicate all of the issues discussed above even in the cases where the populations of particular states are almost entirely comprised of only two racial/ethnic groups.

For further discussions of the problems with this approach to measurement, see Kansas Law Paper (at 23-26), and CEP Comments (at 43-45), TDHCD Brief (at 23-27), IDEA Data Center Disproportionality Guide subpage of the Discipline Disparities page of and slides 97-108 of the University of Maryland workshop.