

**COMMISSION ON EVIDENCE-BASED POLICYMAKING COMMENTS,  
DOCKET ID USBC-2016-0003, QUESTION 16**

Submitted by:

James P. Scanlan  
Attorney at Law  
1529 Wisconsin Ave., NW  
Washington, DC 20007  
jps@jpscanlan.com  
(202) 338-9224

*These comments were submitted through the regulations.gov portal on November 14, 2016, which was the original due date provided with the request for comments. The due date, however, had been extended to December 14, 2016. This version of the comments has some corrections to the document originally submitted. They are listed at the end of the document. Last corrections were made November 23, 2016.*

*Principal issues addressed in Part I include (a) the unsoundness of analyses of demographic and other differences involving outcome rates as a result the failure to consider the ways the measures employed in such analyses tend to be affected by the prevalence of an outcome; and (b) the mistaken belief of federal civil rights law enforcement agencies and a substantial portion of the social science community that relaxing standards and otherwise reducing the frequency of adverse outcomes tend to reduce relative differences in rates of experiencing those outcomes and the proportions disadvantaged groups make up of persons experiencing those outcomes. Part II addresses the unsoundness of analyses of discrimination issues that examine data solely on persons who accepted some outcome or situation. Part III recommends action to address issues discussed in Parts I and II.*

*A summary of the comments may be found in pages 1-8 and their recommendations may be found in pages 45-46. Readers are encouraged to submit comments expressing their agreement or disagreement with the substance of the comments or their recommendations through the regulations.gov portal by the December 14, 2016 due date.*

## Table of Contents

Introduction .....	1
I. The Unsoundness of Analyses of Demographic Differences That Fail to Consider the Effects of the Prevalence of an Outcome on the Measure Employed .....	8
A. Relative Differences in Rates of Experiencing Favorable and Adverse Outcomes .....	14
B. Absolute Differences and the Odds Ratios .....	29
C. The Appraisal of Demographic Differences in Terms of the Proportion a Group Comprises of Persons Potentially Experiencing an Outcome and the Proportion it Comprises of Persons Actually Experiencing the Outcome .....	39
D. Subgroup Effects Issues .....	41
II. The Fundamental Unsoundness of Analyses of Discrimination Issues That Examine Data Solely on Persons Who Accepted Some Outcome or Situation .....	43
III. Recommendations .....	45
Table 1. Illustration of effects on relative differences in pass and fail rates of lowering a cutoff from a point where 80% of AG passes to a point where 95% of AG passes, with proportions DG comprises of persons who pass and of persons who fail (when mean scores differ by approximately half a standard deviation and DG comprises 50% of test takers) .....	20
Table 2. Rates at which white and blacks fall above and below 125%, 100%, and 75% of the poverty line, with measures of differences between rates (2004) .....	22
Table 3. Favorable outcome rates of advantaged group (AG) and disadvantaged group (DG) at four settings with different favorable outcome frequencies, with measures of difference .....	32
Figure 1. Ratios of (1) DG Fail Rate to AG Fail Rate, (2) AG Pass Rate to DG Pass Rate, at Cutoffs Defined by AG Failure Rates .....	22
Figure 2. Absolute differences between rates of AG and DG pass (or fail) rates at various cutoff points defined by AG fail rate .....	31
Figure 3. Ratios of (1) DG fail rate to AG fail rate, (2) AG pass rate to DG pass rate, (3) DG failure odds to AG failure odds .....	31

## INTRODUCTION

Set out below are comments in response to Question 16 of the September 8, 2016 Request for Comments for the Commission on Evidence-Based Policymaking. The question read: “How can data, statistics, results of research, and findings from evaluation, be best used to improve policies and programs?”<sup>1</sup>

The principal subject of these comments, addressed in Part I, involves the facts (a) that virtually all analyses of demographic and other differences involving rates of experiencing favorable or adverse outcomes have been fundamentally unsound as a result of the failure to recognize the ways measures of such differences tend to be affected by the prevalence of an outcome and (b) that, therefore, it is impossible for such analyses to improve policies and programs.<sup>2</sup> One focus of Part I concerns health and healthcare disparities research, a subject to which recently the budget of the National Institutes of Health has devoted over \$2.7 billion per year. Among other problems with such research, a large part of which involves efforts to determine whether health and healthcare disparities are increasing or decreasing over time, it has universally failed to recognize that as health and healthcare generally improve, relative (percentage) differences between the rates at which advantaged and disadvantaged groups experience adverse outcomes (*e.g.*, mortality, non-receipt of appropriate care) tend to increase while relative differences in the rates at which such groups experience the corresponding favorable outcomes (*e.g.*, survival, receipt of appropriate care) tend to decrease. In fact, although more than a decade ago the National Center for Health Statistics recognized that relative differences in adverse health and healthcare outcomes and relative differences in the corresponding favorable health and healthcare outcomes tend to change in opposite directions as health and healthcare generally improve, no other federal agency and no more than a few nongovernmental researchers have recognized that it is even possible for the two relative differences to change in opposite directions as there occur general changes in health and healthcare outcomes.<sup>3</sup>

Thus, researchers analyzing health and healthcare disparities in terms of relative differences in favorable outcomes commonly reach opposite conclusions about directions of changes from researchers examining such disparities in terms of relative differences in the corresponding adverse outcomes. They do so, moreover, without recognizing even the

---

<sup>1</sup> To facilitate consideration of issues raised in documents such as this I include links to referenced materials in electronic copies of the documents. I do not include links with every mention of an item, but attempt to include them often enough to make online items readily available without the reader’s having to search for an earlier link. An electronic copy of this document is available on the [Institutional Correspondence](#) subpage of the [Measuring Health Disparities](#) page of [jpscanlan.com](#). Since the online copy may be corrected or annotated, I include here a [link](#) to the most recent version (which, if corrected or annotated, will so indicate on the first page).

<sup>2</sup> The request for comments and underlying statute are focused on government data. The analyses discussed here are sometime based on government data and sometimes based on other data. But issues regarding soundness of analyses and potential to improve programs are the same regardless of the origin of the data.

<sup>3</sup> I do not want to encumber this introduction with either tabular illustrations or discussions of illustrative numbers. Readers not readily comprehending why relative differences in favorable outcomes and the corresponding adverse outcomes can, or will tend to, change in opposite directions as the prevalence of an outcome changes may wish to refer to Tables 1 and 2 in Section I.A *infra* before proceeding further.

possibility that the relative difference in the opposite outcome could yield a conclusion contrary to that yielded by the relative difference in the outcome they are examining. More important, they do so without thought to the effects of changes in the prevalence of an outcome on the relative difference examined. As of result of this and related failures of understanding regarding the ways measures tend to be affected by the prevalence of an outcome, virtually nothing written about whether health and healthcare disparities are increasing or decreasing over time, or regarding the comparative size of such disparities in one setting compared with another or regarding one outcome compared with another, has had a sound statistical basis. And virtually all of it has been in some manner misleading.

The same holds for analyses of demographic differences regarding matters other than health and healthcare outcomes. For example, no analysis of racial differences regarding things like meeting proficiency standards and graduating from high school has yet recognized that general increases in proficiency and graduation rates will tend to reduce relative racial/ethnic differences in rates of reaching various proficiency levels and graduating while increasing relative differences in rates of failing to reach the proficiency level and failing to graduate. The failure to understand this matter with regard to educational outcomes remains essentially universal notwithstanding that data from the tests underlying such patterns make abundantly clear that lowering a test cutoff, or generally improving test performance, will tend to reduce relative differences in pass rates while increasing relative differences in failure rates.

Another focus of Part I involves the belief underlying federal civil rights law enforcement policies pertaining to fairness in lending, school discipline, criminal justice, and employment that relaxing standards or otherwise reducing the frequency of adverse outcomes will tend to reduce relative racial/ethnic and other demographic differences in rates of experiencing those outcomes. Exactly the opposite is the case. As with lowering a test cutoff, relaxing a standard and thereby reducing the frequency of an adverse outcome, while tending to reduce relative differences in rates of experiencing the corresponding favorable outcome, tends to increase relative differences in rates of experiencing the adverse outcome itself.

In consequence of the failure to understand this matter on the part of federal civil rights enforcement agencies and the research community, for decades countless entities covered by federal civil rights law that have complied with government encouragements to reduce the frequency of adverse outcomes have increased the chances that the government will sue them for discrimination. Such entities include, in recent years, the hundreds or thousands of school districts (and a number of states) that have been relaxing public school discipline standards while mistakenly believing that doing so will tend to reduce relative racial/ethnic and other demographic differences in rates of adverse discipline outcomes like suspension and expulsion. In fact, rarely will a month pass without the prominent reportage of efforts to reduce relative demographic differences in discipline rates by generally relaxing discipline standards, even as jurisdiction after jurisdiction reports that general reductions in discipline rates have been accompanied by increased relative racial/ethnic differences in discipline rates.

Meanwhile, the Department of Justice has been vigorously seeking to reduce relative differences in criminal justice outcomes (or measures that are functions of those relative differences) in Ferguson, Missouri, Baltimore, Maryland, and any number of other jurisdictions,

by modifications of practices in ways that are more likely to increase those differences than reduce them. And reports of a presidential commission and the National Research Council reflect no better an understanding of the connection between the prevalence of adverse criminal justice outcomes and relative differences in experiencing the outcomes than the Department of Justice.

Further, in a range of contexts, entities covered by civil rights laws are required to justify practices with disparate impacts on protected groups and to implement less discriminatory alternatives even to justified practices. The governmental bodies imposing these requirements or monitoring and enforcing compliance with the requirements, as well as those adjudicating disputes as to compliance, do so while unaware that modifications to practices will commonly increase relative differences in one outcome while reducing relative differences in the opposite outcome.

The same misunderstanding reflected in the above-discussed civil rights enforcement activities are involved in perceptions about a wide range of policies directly aimed at mitigating demographic differences in adverse outcomes. For example, programs providing relief against foreclosure on home loans are universally regarded as means of reducing the relative racial/ethnic differences in foreclosure rates. While such programs will tend to reduce relative differences in rates of avoiding foreclosure, they will tend to (in fact will almost certainly) increase relative differences in foreclosure rates.

Recommendations to the Commission pertaining to Part I include (a) that the Commission form a committee to examine the soundness of analyses of differences involving outcome rates and explore methods for improving such analyses; (2) that the Commission recommend to Congress that it establish a permanent body to advise Congress and Executive Branch agencies on the statistical soundness of government-funded and other research; (3) that the Commission recommend that Congress require that all requests for federal funding include a statement as to whether and how the contemplated research will attempt to distinguish between the extent to which changes in the measures employed in research (or the comparative size of such measures in different settings) are functions of the prevalence of an outcome and the extent to which the changes reflect the effects of policies or indicate actual changes in the comparative circumstances of advantaged and disadvantaged groups; (4) that the Commission recommend that Congress take measures to ensure that federal civil rights law enforcement policies, and its own legislative actions, are not based on the mistaken belief that reducing the frequency of adverse outcomes tends to reduce relative demographic differences in rates of experiencing those outcomes; and (5) that the Commission recommend that Congress take measures to ensure that all laws and regulations requiring the monitoring of demographic differences or imposing obligations regarding disparate impacts provide guidance on how such differences and impact are to be measured.

\*\*\*

Recent substantial treatments of issues addressed in Part I, with many examples of the pervasive failure to understand elementary statistical phenomena, may be found in my [“The Mismeasure of Health Disparities,”](#) *Journal of Public Health Management and Practice*

(July/Aug. 2016) (JPHMP commentary),<sup>4</sup> [“Race and Mortality Revisited,”](#) *Society* (July/Aug. 2014), [“The Perverse Enforcement of Fair Lending Laws,”](#) *Mortgage Banking* (May 2014), [“Measuring Health and Healthcare Disparities,”](#) Proceedings of the Federal Committee on Statistical Methodology 2013 Research Conference (March 2014) (FCSM paper), [“The Mismeasure of Discrimination,”](#) Faculty Workshop, University of Kansas School of Law (Sept. 2013) (Kansas Law paper), as well as my amicus curiae [brief](#) in the U.S. Supreme Court in the case of *Texas Department of Housing and Community Development v. The Inclusive Communities Project, Inc.* (Nov. 2014) (TDHCA brief) and my October 8, 2015 [letter](#) to the American Statistical Association (ASA) (October 2015 ASA letter). The most recent extended collection of graphical and tabular illustrations of the pertinent statistical patterns may be found in my quantitative methods seminar [“The Mismeasure of Health Disparities in Massachusetts and Less Affluent Places,”](#) Department of Quantitative Health Sciences, University of Massachusetts Medical School (Nov. 2015) (UMMS seminar). A less recent extended treatment of the issues may be found in my guest editorial in the Spring 2006 issue of the American Statistical Association publication *Chance*, titled [“Can We Actually Measure Health Disparities?”](#) (2006 *Chance* editorial), to which I will make a number of references below.<sup>5</sup>

Recent, more succinct treatments of key statistical issues, with a focus on federal law enforcement anomalies, may be found in my [“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), [“Things DoJ doesn’t know about racial disparities in Ferguson,”](#) *The Hill* (Feb. 22, 2016), [“Things government doesn’t know about racial disparities,”](#) *The Hill* (Jan. 28, 2014), [“The Paradox of Lowering Standards,”](#) *Baltimore Sun* (Aug. 5, 2013), and [“Misunderstanding of Statistics Leads to Misguided Law Enforcement Policies,”](#) *Amstat News* (Dec. 2012). A recent, somewhat more extended treatment of the law enforcement anomalies, with a focus on misunderstandings reflected in President Barack Obama’s July 7, 2016 speech on racial/ethnic differences in criminal justice outcomes, may be found in my [“Things the President Doesn’t Know About Racial Disparities,”](#) *Federalist Society Blog* (Aug. 5, 2016). A recent succinct treatment of the pattern whereby relative differences in adverse outcomes tends to be larger, while relative differences in the corresponding favorable outcomes tend to smaller, in settings where the adverse outcomes are comparatively uncommon (including nations like Norway and Sweden and states like Minnesota) than in settings where adverse outcomes are comparatively common may be found in my [“It’s easy to misunderstand gaps and mistake good fortune for a crisis,”](#) *Minneapolis Star Tribune* (Feb. 8, 2014).<sup>6</sup>

---

<sup>4</sup> These comments do not attempt to employ a standard or consistent citation format. They simply employ the formatting that is most useful or convenient in the circumstances, in some cases simply copying the citations used in other works.

<sup>5</sup> The recent extended treatments all discuss a measure of differences between outcome that is theoretically unaffected by the prevalence of the outcome, while the 2006 *Chance* editorial does not.

<sup>6</sup> The patterns discussed in the *Star Tribune* commentary with regard to Norway, Sweden, and Minnesota are addressed with regard to Massachusetts in the UMMS seminar. See the UMMS seminar [abstract](#). See also the discussion of patterns of discipline disparities in Massachusetts in the letter to the [Boston Lawyers’ Committee for Civil Rights and Economic Justice](#) (Nov. 12, 2015), which provides a narrative explanation of the data in Tables B1 and B2 (slides 67 and 68) of the UMMS seminar.

A recent, brief treatment of the uncertainties facing entities with obligations regarding the disparate impacts of practices as a result of the near universal failure to understand the ways measures tend to be affected by the prevalence of an outcome may be found in my "[Is HUD's Disparate Impact Rule Unconstitutionally Vague?](#)," *American Banker* (Nov. 10, 2014). A recent, more comprehensive treatment of the subject may be found in "[Is the Disparate Impact Doctrine Unconstitutionally Vague?](#)," *Federalist Society Blog* (May 6, 2016). See also Part II the TDHCD brief.

\*\*\*

As reflected in the October 2015 ASA [letter](#), on October 8, 2015, I requested that ASA take certain actions to reform analyses of demographic differences and to explain to arms of the federal government that reducing the frequency of an adverse outcome will tend to increase, rather than reduce, relative racial and other demographic differences in rates of experiencing the outcome. The matter was referred to the organization's Science and Public Affairs Advisory Committee, which subsequently sought additional information regarding situations where the government encouraged the relaxing of standards to reduce relative differences in adverse outcomes. I responded by memoranda of [December 14, 2015](#) and [January 5, 2016](#). In a [letter](#) to ASA of July 25, 2016, in addition to informing the organization of developments since the October 2015 letter, I urged ASA to explain to President Barack Obama that understandings about effects of modifications of practices on measures of racial/ethnic differences in criminal justice outcomes reflected in the President's July 7, 2016 speech on criminal justice disparities were incorrect.

In September 2016, the ASA informed me that the Science and Public Affairs Advisory Committee believed that I was effectively highlighting the issues raised in my communications to the organization and did not see an additional role ASA could play in disseminating information on such issues. But, whereas ASA has chosen not to take affirmative steps to promote the understanding of such issues, as a result of the organization's consideration of my October 2015 ASA letter and other communications by its Science and Public Affairs Subcommittee, the organization should be in an excellent position to advise the Commission regarding the validity of the points made in Part I of these comments. That holds both for the broader points in the comments and for the explanation that the government's belief that generally reducing adverse lending, school discipline, criminal justice, and employment outcomes will tend to reduce relative differences in rates of experiencing those outcomes is not merely incorrect, but is the opposite of reality.<sup>7</sup>

---

<sup>7</sup> Explanations of why reducing the prevalence of an outcome tends to increase relative differences in experiencing the outcome (while reducing relative differences in the corresponding outcome) in ASA publications include the December 2012 *Amstat News* column and Spring 2006 *Chance* guest editorial previously mention and the article "[Divining Difference](#)," *Chance* (Fall 1994).

By letter of March 28, 2016, I requested the same action of the Population Association of American (PAA) and Association of Population Centers (APC) that I had requested of the ASA. At the end of April, the organizations advised that they were taking none of the requested actions because they did not prescribe statistical methods to the government or to their members. As with ASA, however, the leadership of these organizations should be able to advise the Commission as to the validity of the points I make in Part I of these comments.

In seeking the views of any person or organization of presumptive expertise, however, the Commission should be mindful that, so far as the published record reveals, there have been no recognitions that relative differences in experiencing an outcome and relative differences in avoiding the outcome tend to change in opposite directions as the prevalence of the outcome changes (or even that reducing the prevalence of an adverse outcome tends to increase relative differences in rates of experiencing it) other than in response to my explanations of these patterns.<sup>8</sup> And, for example, the belief that reducing adverse outcomes – whether regarding health and healthcare, lending, school discipline, criminal justice, or employment – will tend to reduce relative differences in experiencing those outcomes is widespread within the scientific community, while the understanding that it will tend to increase those differences is essentially non-existent (National Center for Health Statistics excepted).<sup>9</sup> Thus, when first asked, an experienced statistician or other data analyst might well say that a statement that lowering test cutoffs tends to increase relative differences in failure rates or that reducing poverty tends to increase relative differences in poverty rates is plainly false; but on reflection, and simply viewing the first two tables of these comments, the same person would almost certainly say that the statement is plainly true. Thus, while I encourage the Commission to seek input from the scientific community on the issues raised in these comments, I urge it to ensure that the persons providing such input consider the issues in depth.

---

<sup>8</sup> I emphasize the published record. Given the elementary nature of the patterns I describe and that they are obvious in so many types of data, I assume that countless individuals with and without statistical training have independently recognized the patterns even if they have not had occasion to publish such recognitions. See discussion of the work of Andrew Ho in Section I.B regarding the recognition of the pattern by which absolute differences tend to be affected by the prevalence of an outcome that might be deemed implicit recognition of the pattern by which the two relative differences tend to be affected by the prevalence of the outcome. Further, the patterns I describe as to relative differences in both outcomes are implicit in the widely understood pattern by which relative differences in favorable outcomes tend to be affected by the prevalence of an outcome. See October 2015 ASA letter at 14.

<sup>9</sup> A reflection of the current understanding of this subject in the social science community with respect to school discipline issues may be found in an article in the Winter 2017 issue of *Education Next* titled “[What Do We Know About School Discipline Reform](#).” Calling for research into whether modifications to practices that generally reduce discipline rates will in fact reduce relative racial/ethnic differences in discipline rates, the article cited one recent study showing that after substantial reductions in suspensions, relative racial/ethnic differences in discipline rates “continue” to be high. While referencing an instance where the conventional understanding was not borne out, the article showed no awareness whatever of the reasons to expect general reductions in discipline rates to increase relative racial/ethnic differences in suspension rates or of the numerous situations across the country in which such increases are in fact being observed. Current titles reflecting the near universal expectation that reductions in the prevalence of an outcome should reduce relative differences in experiencing it (or a measure that is a function thereof) are commonplace. See, among many others, “[Despite Progress, Racial Disparities Persist](#),” Sentencing Project (Aug. 19, 2016) and “[Huge racial disparities persist despite slow infant mortality drop](#),” USA Today (Mar. 9, 2016).



Finally, I note that, while most of my longer recent papers and all of my workshops and conference presentations since 2008 discuss a method for appraising the difference between the circumstances of advantaged and disadvantaged groups reflected by their outcome rates (sometimes characterized as the strength of the forces causing outcome rates to differ) that is unaffected by the prevalence of an outcome, I give little attention to such measure here. The purpose of these comments is to cause the Commission to address the failings of standard analyses of differences in outcome rates and to explore methods for effectively measuring such differences. The measure I have discussed is merely one of a number of measures the Commission may consider.<sup>10</sup> Further, attention to my proposed measure has on occasion diverted attention from the crucial issue of whether it is possible to analyze demographic differences in ways that can inform policy without consideration of the ways the measures employed tend to be affected by the prevalence of an outcome.<sup>11</sup> Thus, these comments can better serve their purpose with only limited attention to the measure that I have discussed at length elsewhere.

\*\*\*

Part II of the comments addresses the fundamental unsoundness of analyses of discrimination issues that are based on the examination of data solely on persons who accepted some outcome or situation and thus fail to consider the entire universe of persons seeking the most desirable outcome or situation. Almost all monetary recoveries in discrimination cases that have approached or exceeded \$100 million have involved such analyses. These include the recent settlements of lending discrimination cases brought by the Department of Justice against Countrywide Financial Corporation and Wells Fargo Bank, with total recoveries exceeding half a billion dollars, that are discussed in the December 2012 *Amstat News* column and several other of the above references. This subject is treated in Section F (at 32-35) of the Kansas Law [paper](#) and Section I.C (at 27-30) of the TDHCD [brief](#).

The subject is treated fairly succinctly with reference to fair lending issues in “[Fair Lending Studies Paint Incomplete Picture](#),” *American Banker* (April 24, 2013) and on the final page of “[The Perverse Enforcement of Fair Lending Laws](#),” *Mortgage Banking* (May 2014). Older treatments of varying length involving employment discrimination issues may be found in my “[Illusions of Job Segregation](#)” *Public Interest* (Fall 1988), “[Multimillion-Dollar Settlements May Cause Employers to Avoid Hiring Women and Minorities for Less Desirable Jobs to Improve the Statistical Picture](#),” *National Law Journal* (Mar. 27, 1995), and several other items referenced on the [Employment Discrimination](#) page of [jpscanlan.com](#).

---

<sup>10</sup> See “Race and Mortality Revisited” at 337. See also discussion of the odds ratio in note 49 *infra*.

<sup>11</sup> See (a) Mackenbach JP. [Response to Scanlan](#). *BMJ* (July 14, 2016), and (b) Penman-Aguilar, Talih M, Moonesinghe R, Huang M. “[Response to Scanlan Concerning: Measurement of Health Disparities, Health Inequities, and Social Determinants of Health to Support the Advancement of Health Equity](#).” *J Public Health Management Practice* 2016;22(6), 614-615 [litr]. While focusing on the measure I suggest as theoretically unaffected by the prevalence of an outcome, these responses leave entirely unaddressed whether health and healthcare disparities research can be of any value without considering the ways the measures employed tend to be affected by the prevalence of an outcome.

Analyses that fail to examine the entire universe at issue are also involved in the cases of *Bank of America Corp. et al. v. City of Miami* and *Wells Fargo & Co. v. City of Miami*, Sup. Ct. Nos. 15-1111 and 15-1112, which were argued before the Supreme Court on November 8, 2016 (though the issue before Court do not involve the merits of the claims).<sup>12</sup>

This subject is given brief attention at page 32 of the October 2015 ASA letter. But I would not regard the organization's consideration of my letter as putting it in a position to have a sound understanding of the subject. The organization, of course, does have the expertise to provide informed views on the issue should the Commission seek such views.<sup>13</sup>

Recommendations to the Commission regarding issue addressed in Part II are similar in to the above-mentioned recommendations regarding issues addressed in Part I.

Part III discusses the recommendations to the Commission regarding the issues addressed in Parts I and II.

## **I. The Unsoundness of Analyses of Demographic Differences That Fail to Consider the Effects of the Prevalence of an Outcome on the Measure Employed**

There are four standard measures by which observers commonly quantify differences between rates at which advantaged and disadvantaged groups experience favorable or adverse outcomes: (1) relative (percentage) differences between rates of experiencing the outcome; (2) relative differences between rates of avoiding the outcome (i.e., experiencing the opposite outcome); (3) absolute (percentage point) differences between the outcome rates; and (4) odds ratios. None of these measures provides a sound basis for quantifying the differences in the circumstances of advantaged and disadvantaged groups reflected by their outcome rates because, for reasons related to the shapes of underlying risk distributions, each measure tends to be systematically affected by the prevalence of an outcome.<sup>14</sup>

---

<sup>12</sup> Claims in these cases also implicate the subjects of Part I with regard to the failures to recognize that (a) the less the incentive for loan officers to issue subprime loans, the larger (not smaller) will tend to be relative racial/ethnic differences in receipt of subprime loans; (b) concentrations of foreclosures in disadvantaged neighborhoods will vary inversely with the frequency of foreclosures; (c) relative differences in assignment to subprime status and foreclosures will tend to be larger, while relative differences in avoiding those outcomes will tend to be smaller, among higher-income than lower-income borrowers. See, e.g., the 2014 *Mortgage Banking* article.

<sup>13</sup> The letter to the Population Association of America and the Association of Populations Centers did not mention this subject at all. As with ASA, however, the organizations should be able to form views on the subject if requested by the Commission.

<sup>14</sup> As will be discussed in Sections A and C, observers often appraise demographic differences in terms, not of differences between outcome rates, but of differences between the proportion a group comprises of persons potentially experiencing an outcome and proportion it comprises of persons actually experiencing the outcome. It is for that reason that the Introduction refers to "differences involving outcome rates" rather than "differences between outcome rates." These comments will give attention both to the relationship between differences between outcome rates and differences between the two referenced proportions and to the impossibility of soundly analyzing demographic differences based solely on information regarding the two proportions. But, for simplicity, issues regarding the two proportions are given minimal attention in introductory material.

By way of broad summary, as an outcome changes in prevalence, the relative difference in rates of experiencing it and the relative difference in rates of avoiding it tend to change in opposite directions. More specifically, the relative difference in the decreasing aspect of the outcome tends to increase while the relative difference in the corresponding increasing aspect of the outcome tends to decrease. As the prevalence of an outcome changes, absolute differences between rates and differences measured by odds ratios tend also to change as the prevalence of an outcome changes, though in a more complicated way than the two relative differences. Roughly, as an outcome goes from being rare to being common, the absolute difference tends to increase; as an outcome goes from being common to being very common, the absolute difference tends to decrease. As the frequency of an outcome changes, the absolute difference tends to change in the same direction as the smaller relative difference. Since persons relying on relative differences to appraise demographic differences typically examine the larger of the two relative differences, such persons tend to systematically reach opposite conclusions about directions of changes over time from persons relying on absolute differences. As the prevalence of an outcome changes, the difference measured by the odds ratio tends to change in the opposite direction of the absolute difference.

As the prevalence of an outcome changes, all measures may change in the same direction. In that case, one may infer that there occurred an actual change in the difference between the circumstances of two groups reflected by their outcome rates. But whenever a relative difference and the absolute difference change in different directions, the other relative difference will necessarily have changed in the opposite direction of the first relative difference and the same direction as the absolute difference.

For a variety of reasons, including actual changes in the strength of the forces causing outcome rates to differ (or actual differences in the strength of those forces in different settings),<sup>15</sup> one may observe many departures from these patterns. But it is impossible to evaluate the efficacy of policies aimed at mitigating differences in the circumstances of advantaged and disadvantaged groups (or to draw inference about underlying processes) without understanding the above patterns and attempting to distinguish between the effects of changes in measure that are functions of changes in the prevalence of an outcome and changes that are functions of other factors. It is also impossible to determine whether a particular measure indicates a large or small difference between the circumstances of advantaged and disadvantaged groups without understanding the effects of the prevalence of an outcome on the measure.

Further, the patterns I describe may provide useful benchmarks for identifying flawed interpretations of data (and potentially for divining a means of measuring demographic differences unaffected by the prevalence of an outcome). But with regard to the fundamental

---

<sup>15</sup> Other factors include sampling variation, irregularities in the underlying distributions, factors related to the approaching of an irreducible minimum regarding outcomes like infant mortality. They also include the possibility that the underlying distributions may substantially depart from the normal distribution, among other reasons, because the universe being examined is comprised of truncated portions of normal distributions. See "Race and Mortality Revisited" and the materials it references (at 337) regarding factors creating potential problems with the measure discussed there. The same factors affect the likelihood of observing the described patterns in particular situations.

unsoundness of the standard approaches to analyzing demographic differences, the degree to which the measures in fact tend to be affected by the prevalence of an outcome comports with my descriptions is unimportant, as is the precise utility of the measure I suggest in various places is theoretically unaffected by the prevalence of an outcome. For as long as a measure tends to be affected by the prevalence of an outcome in any manner, it is not possible to employ the measure for a useful purpose without attempting to divine the role of the prevalence of the outcome in the setting examined. Indeed, the mere possibility for the patterns shown in the tables of these comments to exist in nature raises the same interpretative issue as the fact that such patterns will commonly, if not almost invariably, be found whenever there exist substantial differences in the prevalence of an outcome in the settings being compared.<sup>16</sup>

Yet, with few or no exceptions, persons analyzing demographic differences have done so without even recognizing that the measures employed tend to change as the frequency of an outcome changes and hence have never attempted to divine the role of the prevalence of an outcome in the settings examined. Thus, almost all such analyses have been statistically unsound and in some manner misleading. Even when a study may have correctly identified such things as the change in the direction of a demographic disparity, it will have been misleading by implying that the measure employed is effectively quantifying either the difference in the circumstances of two groups reflected by their outcome rates or the size of the change in such difference. Even a study that merely reports the size of a disparity is commonly misleading by suggesting that the measure employed is effectively quantifying the disparity between the circumstances of the groups being compared.<sup>17</sup>

Section A addresses the patterns by which relative differences in a favorable outcome and relative differences in the corresponding adverse outcome tend to be affected by the prevalence of an outcome and some implications of the failure to understand those patterns, while discussing other measures only to the extent necessary to make particular points. It gives particular attention to interpretations of health and healthcare disparities based on relative differences in a favorable or adverse outcomes and the 2004-05 recommendation of the National Center for Health Statistics regarding which relative difference to rely on, as well as the recent reversal of that recommendation and the repudiation of a decade of research that relied on the earlier recommendation. The section also gives particular attention to federal law enforcement policies

---

<sup>16</sup> See the October 2015 ASA letter (at 27) regarding the interaction of other factors with the prevalence-related patterns described here, including discussion of why the patterns will be having an influence even when they are not specifically observed.

<sup>17</sup> Studies that attempt to determine the role of a particular factor in an observed disparity may not suffer from same problems as studies that attempt to examine changes over time or otherwise to compare the size of disparities in difference settings, at least so far as the primary purpose of the former studies is concerned. But the issues addressed here do affect the quantification of the observed disparity before and after adjustment for the factor and thus the quantifications of the part of the difference that is accounted for by the factor. That is so even though, say, in a situation where when the rates at which an advantaged group and a disadvantaged group experience an adverse outcome are 10% and 20% and adjustment for a factor would reduce the latter rate to 15%, one would observe the same 50% reduction in both relative differences and in the absolute difference. See [Comment on Lynch JECH 2006](#) (2006) and [Second Comment on Lynch JECH 2006](#) (2009). But one would not observe that same reduction for the difference measured by the odds ratio or the measure discussed in "Race and Mortality Revisited."

based on the mistaken belief that reducing the frequency of an adverse outcome will tend to reduce relative racial/ethnic and other demographic differences in rates of experiencing it.

Section B discusses the patterns by which absolute differences and odds ratios tend to be affected by the prevalence of an outcome and some of the implications of the failure to understand those patterns or the relationship of those patterns to the patterns by which the two relative differences tend to be affected by the prevalence of an outcome. The section gives particular attention to the analyses of healthcare disparities and public school proficiency disparities in terms of absolute differences between rates without consideration of the ways improvements in healthcare or education will tend to increase absolute differences regarding uncommon outcomes and reduce absolute differences regarding common outcomes.

Section C discusses problems with analyses of demographic differences based on the proportion a group comprises of persons potentially experiencing an outcome and the proportion it comprises of persons actually experiencing the outcome that go beyond the problems involved in standard analyses of differences between outcome rates. The section explains why a demographic difference should never be analyzed on the basis of comparisons of the two proportions.

Section D discusses problems with standard analyses of subgroup effects arising from the failure to understand the patterns described here and the failure to understand the illogic of an expectation that a factor will tend to have the same relative effect on an outcome in situations involving different baseline rates for the outcome.

The reader will find considerable redundancy in the discussion in Sections A through D, especially with regard to relative measures. To some degree the redundancy occurs because many points are necessarily implied in other points. For example, the statement that reducing the prevalence of an outcome tends to increase relative differences in the outcome, by necessarily implying that increasing the prevalence of an outcome tends to reduce relative differences in experiencing the outcome, necessarily also implies that reducing the prevalence of an outcome tends to reduce relative differences in the opposite outcome. The statement that as an outcome changes in prevalence the group with the lower baseline rate tends to experience a larger proportionate change in the rate than the other group necessarily means that the other group will tend to experience the larger proportionate change in the opposite outcome, since the other group has the lower baseline rate for the opposite outcome. See October 2015 ASA letter (at 10 n.14). The statement that relative racial differences in some adverse outcome like rejection of a loan application or poor self-rated health tend to be greater, while relative racial differences in the corresponding favorable outcomes tend to be smaller, among higher-income than lower-income groups means the same thing as the statement that having high income tends to reduce the adverse outcome proportionately more for whites than blacks while increasing the favorable outcome proportionately more for blacks than whites. In fact, all statements about failure to understand patterns of relative differences are effectively statements about misperceptions regarding subgroup effects. And, of course, the statements made above about higher- and lower-income groups effectively illustrate the fallacy of essentially all perceptions about racial/ethnic or other differences in favorable or adverse outcomes within advantaged subpopulations or the effects of being in an advantaged subpopulation on different racial/ethnic groups.

Nevertheless, I believe it useful to make various points that are already implied in other points in order (a) to make the points clearer to readers to whom the relationship between the points may not be obvious, (b) to illustrate how commonly the patterns I describe are found in reality, and (c) to illustrate how vast is the waste of resources devoted to study of demographic differences, and the misleading reportage of the results thereof, in consequence of the failure to understand the patterns described in this part.

The discussion that follows contains a number of tabular or graphical illustrations, as do the more extended treatments of this subject listed above. Larger collections of graphical and tabular illustrations, which include many examples from actual studies, may be found in the methods workshops given at American universities between 2012 and 2015,<sup>18</sup> and the conference presentations given in Europe and North American between 2001 and 2011.<sup>19</sup>

Letters to institutions and organizations written since 2009 provide numerous examples of misinterpretations of data by entities believed to have considerable expertise in the analyses of demographic differences and suggest the universality of the unsoundness of such analyses even among organizations and institutions whose missions principally involve the analysis of such differences.<sup>20</sup> Other examples may be found in online comments to medical and health policy

---

<sup>18</sup> In addition to the November 2015 UMMS seminar referenced *supra*, see [“The Mismeasure of Discrimination,”](#) Center for Demographic and Social Analysis, University of California, Irvine (Jan. 20, 2015); [“The Mismeasure of Demographic Differences in Outcome Rates”](#) Public Sociology Association of George Mason University (Oct. 18, 2014); [“Rethinking the Measurement of Demographic Differences in Outcome Rates,”](#) Maryland Population Research Center of the University of Maryland (Oct. 10, 2014); [“The Mismeasure of Association: The Unsoundness of the Rate Ratio and Other Measures That Are Affected by the Prevalence of an Outcome,”](#) Minnesota Population Center and Division of Epidemiology and Community Health of the School of Public Health of the University of Minnesota (Sept. 5, 2014); [“The Mismeasure of Group Differences in the Law and the Social and Medical Sciences,”](#) Institute for Quantitative Social Science at Harvard University (Oct. 17, 2012); [“The Mismeasure of Group Differences in the Law and the Social and Medical Sciences,”](#) Department of Mathematics and Statistics of American University (Sept. 25, 2012).

<sup>19</sup> There were twenty-two such presentations given in eight countries (mainly between 2006 and 2011). Links to them are collected [here](#). Like the methods workshops, many presentations bear much similarity to one another. Some, however, give particular attention to misunderstandings pertinent to a particular country or region, as in the case of the 2006 British Society for Populations Studies paper [“The Misinterpretation of Health Inequalities in the United Kingdom”](#) ([presentation](#)) and the 5th Nordic Health Promotion Research Conference (2006) presentation [“The Misinterpretation of Health Inequalities in Nordic Countries”](#) ([abstract](#)), which is discussed in the above-referenced 2014 *Minneapolis Star Tribune* commentary. The 16th Nordic Demographic Symposium (2008) presentation [“Measures of Health Inequalities that are Unaffected by the Prevalence of an Outcome”](#) uses illustrations from studies of demographic differences in health and healthcare outcomes pertaining to Finland, where the conference was held. The 2011 International Conference on Health Policy Statistics presentation [“Perverse Perceptions of the Impact of Pay for Performance on Healthcare Disparities”](#) ([abstract](#)) focuses on misperceptions about the effects of pay-for-performance programs on healthcare disparities, a subject later addressed more fully in “Race and Mortality Revisited” and the FCSM paper and a number of the methods workshops. The 2011 Joint Statistical Meetings presentation [“Interpreting Differential Effects in Light of Fundamental Statistical Tendencies”](#) ([abstract](#)) focused on perceptions about subgroup effects. Only studies from 2008 on, commencing with the 7th International Conference on Health Policy Statistics presentation [“Can We Actually Measure Health Disparities?”](#) ([abstract](#)), discuss a method for measuring demographic differences unaffected by the prevalence of an outcome.

<sup>20</sup> Recipient of such letters include entities whose activities involves the analyses of demographic differences or the oversight of entities analyzing demographic differences as well as entities who are affected by analyses of

journal since 2005.<sup>21</sup> Further examples of misinterpretations of demographic differences and explanations as to certain nuances of patterns by which measures tend to be affected by the prevalence of an outcome may be found on the pages and subpages of [jpscanlan.com](http://jpscanlan.com) devoted to measurement issues.<sup>22</sup>

Many examples may also be found in older articles addressing issues that were then topical or had been for some time.<sup>23</sup> For example, the 1987 *Plain Dealer* article "[‘Feminization of Poverty’ is Misunderstood](#)" discusses the failure to recognize that reductions in

---

demographic differences by other entities. Recipients include [Oklahoma City School District](#) (Sept. 20, 2016), [Antioch Unified School District](#) (Sept. 9, 2016), [American Statistical Association II](#) (July 25, 2016), [Federal Judicial Center](#) (July 7, 2016), [University of Oregon Institute on Violence and Destructive Behavior and University of Oregon Law School Center for Dispute Resolution II](#) (July 5, 2016), [University of Oregon Institute on Violence and Destructive Behavior and University of Oregon Law School Center for Dispute Resolution](#) (July 3, 2016), [New York City Center for Innovation through Data Intelligence](#) (June 6, 2016), [Consortium of Social Science Associations](#) (Apr. 6, 2016), [Population Association of America and Association of Population Centers](#) (Mar. 29, 2016), [Council of Economic Advisers](#) (Mar. 16, 2016), [City of Madison, Wisconsin](#) (Mar. 12, 2016), [Stanford Center on Poverty and Inequality](#) (Mar. 8, 2016), [City of Boulder, Colorado](#) (Mar. 5, 2016), [Houston Independent School District](#) (Jan. 5, 2016), [Boston Lawyers’ Committee for Civil Rights and Economic Justice](#) (Nov. 12, 2015), [House Judiciary Committee](#) (Oct. 19, 2015), [American Statistical Association](#) (Oct. 8, 2015), [Chief Data Scientist of White House OSTP](#) (Sept. 8, 2015), [McKinney, Texas Independent School District](#) (Aug. 31, 2015), [Department of Health and Human Services and Department of Education](#) (Aug. 24, 2015), [Agency for Healthcare Research and Quality](#) (July 1, 2015), [City of Minneapolis, Minnesota](#) (June 8, 2015), [Texas Appleseed](#) (Apr. 7, 2015), [Senate Committee on Health, Education, Labor and Pensions](#) (Mar. 20, 2015), [United States Department of Justice and City of Ferguson, Missouri](#) (Mar. 9, 2015), [Vermont Senate Committee on Education](#) (Feb. 26, 2015), [Portland, Oregon Board of Education](#) (Feb. 25, 2015), [Wisconsin Council on Families and Children’s Race to Equity Project](#) (Dec. 23, 2014), [Financial Markets and Community Investment Program, Government Accountability Office](#) (Sept. 9, 2014), [Education Law Center](#) (Aug. 14, 2014), [IDEA Data Center](#) (Aug. 11, 2014), [Institute of Medicine II](#) (May 28, 2014), [Annie E. Casey Foundation](#) (May 13, 2014), [Education Trust](#) (April 30, 2014), [Investigations and Oversight Subcommittee of House Finance Committee](#) (Dec. 4, 2013), [Mailman School of Public Health of Columbia University](#) (May 24, 2013), [Senate Committee on Health, Education, Labor and Pensions](#) (Apr. 1, 2013), [Federal Reserve Board](#) (March 4, 2013), [Harvard University et al.](#) (Oct. 26, 2012), [Harvard University](#) (Oct. 9, 2012), [United States Department of Justice](#) (Apr. 23, 2012), [United States Department of Education](#) (Apr. 18, 2012), [The Commonwealth Fund](#) (June 1, 2010), [Institute of Medicine](#) (June 1, 2010), [National Quality Forum](#) (Oct. 22, 2009), [Robert Wood Johnson Foundation](#) (Apr. 8, 2009).

<sup>21</sup> There are approximately 150 such comments, mainly involving interpretations of data on health and healthcare disparities but in some cases (particularly BMJ comments) involving perceptions about subgroup effects. Links to most are available [here](#).

<sup>22</sup> The principal measurement pages are [Measuring Health Disparities](#), [Scanlan’s Rule](#), [Mortality and Survival](#), [Statistical Reasoning](#), [Immunization Disparities](#), [Educational Disparities](#), [Disparate Impact](#), [Discipline Disparities](#), [Lending Disparities](#), [Employment Discrimination](#), [Feminization of Poverty](#). The pages have close to 100 subpages. Brief summaries of the pages and most of their subpages may be found on the home page of [jpscanlan.com](http://jpscanlan.com).

<sup>23</sup> In addition to the 2006 *Chance* editorial and 1994 *Chance* article mentioned *supra*, see "[‘Race and Mortality,’ Society](#) (Jan./Feb. 2000) (reprinted in *Current*, Feb. 2000), "[‘Mired in Numbers,’ Legal Times](#) (Oct. 12, 1996), "[‘When Statistics Lie’](#)" (*Legal Times*, Jan. 1 1996), "[‘Getting it Straight When Statistics Can Lie,’](#)" *Legal Times* ( June 23, 1993), "[‘Bias Data Can Make the Good Look Bad,’](#)" *American Banker* (Apr. 27, 1992), "[‘Comment on ‘McLanahan, Sorensen, and Watson’s ‘Sex Differences in Poverty, 1950-1980’,’](#)" *Signs* (Winter 1991), "[‘An Issue of Numbers,’](#)" *National Law Journal* (Mar. 5, 1990), "[‘The ‘Feminization of Poverty’ is Misunderstood,’](#)" *Plain Dealer* (Nov 11, 1987) (reprinted in *Current*, May 1988, and *Annual Editions: Social Problems 1988/89*, Dushkin 1988).

poverty will tend to increase relative differences between the poverty rates of female-headed families and other units and cause poverty to become more feminized (while increases in poverty will tend to have the opposite effect). The 1990 *National Law Journal* article “[An Issue of Numbers](#),” explains that the high proportion blacks comprised of persons disqualified from intercollegiate athletics by National Collegiate Athletic Association academic standards was a function of the leniency rather than the stringency of the standards. The 1993 *Legal Times* article “[Getting it Straight When Statistics Can Lie](#)” discusses the failure to recognize that improving the performance of all employees (which had been suggested as a means of reducing large relative racial differences in terminations for inadequate performance) would tend to increase relative racial differences in such terminations or that employers that strive for fairness by providing procedural safeguards against arbitrary termination will tend to show larger relative differences in termination rates than other employers. It also discusses a court of appeals decision premised on the belief that a large relative difference in failure to meet a performance standard was a consequence of the stringency, rather the leniency, of the standard. The 1996 *Legal Times* article “[When Statistics Lie](#)” discusses a putative class action based on a study that ranked lenders according to the size of relative racial difference in mortgage rejection rates without recognizing that more lenient lending criteria tended to be associated with larger relative differences in rejection rates. The 1996 *Legal Times* article “[Mired in Numbers](#)” explains why relaxing a three-strikes law will tend to increase the proportion blacks make up of persons affected by the law. The 1991 *Public Interest* article “[The Perils of Provocative Statistics](#)” and the 2000 *Society* article “[Race and Mortality](#)” provide a wide range of examples.<sup>24</sup> The failures of understanding reflected in these examples are just as pervasive today as they were when the articles were published.

#### **A. Relative Differences in Rates of Experiencing Favorable and Adverse Outcomes**

The section discusses the pattern 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 such groups avoid the outcome (i.e., experience the opposite outcome).<sup>25</sup> It also discusses some of the implications of the failure to understand the pattern in analyses of demographic differences.

Discussion of patterns by which absolute differences tend to be affected by the prevalence of an outcome will be largely deferred to the next section. But it would be useful for readers to understand at this point that all cases where a relative difference between rates at which two groups experience an outcome changes in a different direction from the absolute difference between those rates involve situations (a) where the relative difference between rates of

---

<sup>24</sup> References to medical and health policy journal articles underlying many of the statements about health outcomes in “Race and Mortality,” and tables based on data from those articles, may be found in my unpublished paper “[The Relationship Between Declining Mortality and Increasing Racial and Socioeconomic Disparities in Mortality](#)” (1992).

<sup>25</sup> A more precise description of the pattern would state, rather than “the rarer an outcome,” “the more the outcome is restricted toward either end of the overall distribution.” But I have characterized the pattern in the manner done in the text above for some time and those discussing it have not been confused by the usage. Thus, I am not at this time inclined to depart from the usage in the text.



experiencing the opposite outcome has changed in the opposite direction of the relative differences between rates of experiencing the first outcome and (b) where the relative difference between rates of experiencing the increasing outcome has decreased while the relative difference between rates of experiencing the decreasing outcome has increased.<sup>26</sup> The same holds for situations where the relative difference is larger in one setting while the absolute difference is larger in the other setting.<sup>27</sup> That is, the relative difference for one outcome will be larger in the setting where that outcome is less common while the relative difference for the opposite outcome will be larger in the other setting (which is the setting where the opposite outcome is less common). Thus, all situations where attention has recently been given to the fact that a relative difference and absolute difference show different patterns as to the comparative size of some disparity involve situations where the relative difference in rates of experiencing an outcome is larger, while the relative difference in avoiding the outcome is smaller, where the outcome is less common.

Tables 1 and 2 illustrate the pattern whereby the rarer an outcome the greater tend to be the relative differences in experiencing it and the smaller tends to be the relative differences in avoiding it by means of test score and income data. The tables will show how lowering a test cutoff will tend to increase relative differences in failure rates while reducing relative differences in pass rates and how reducing poverty tends to increase relative differences in poverty rates while reducing relative differences in rates of avoiding poverty. But the patterns that will be shown in the tables may be found in virtually any data showing the proportions of two groups

---

<sup>26</sup> I have often noted (usually in discussions of perceptions about value judgments involved in choosing to appraise a demographic disparity in terms of the relative difference the observer happens to be examining or the absolute difference) that anytime it is mentioned that a relative difference and the absolute difference have changed in opposite directions, the unmentioned relative difference will necessarily have changed in the opposite direction of the mentioned relative difference and the same direction as the absolute difference. See, *e.g.*, "Race and Mortality Revisited" (at 335) and "The Mismeasure of Health Disparities" (at 415). But, at least partly because I think many readers would regard the point to be obvious, I have not often explained why that is so. My [Comment on Boscoe BMJ 2015](#) (2016) provides a rather inefficient explanation. The point may be somewhat more efficiently explained in the following paragraph.

In order for a relative difference and the absolute difference for an outcome to change in opposite directions, the rates for both groups must change in the same direction, and the group with the lower baseline rate must experience the larger relative change while the other group experiences the larger absolute change. Any other scenario would cause the relative difference and the absolute difference to change in the same direction. That includes the scenario in which the group with the lower baseline rate experiences the larger absolute change, which necessarily means that it also experiences the larger relative change and that therefore the relative and absolute difference will change in the same direction. The fact that the group with the higher baseline rate for the outcome experiences the larger absolute change in the outcome means that it also experiences the larger absolute change in the opposite outcome. Since that group has the lower baseline rate for the opposite outcome, the fact that it experiences the larger absolute change in the opposite outcome means that it also experiences the larger relative change in the opposite outcome. Thus, the absolute difference and the relative difference for that outcome must change in the same direction.

<sup>27</sup> I sometimes use the phrase "differentiated other than temporally" to refer to comparisons of settings that do not involve changes over time (i.e., comparisons of one place or subpopulation with another or one condition with another). But all involve the same statistical phenomenon. And regardless of the nature of the situations being compared, the extent to which the described patterns are observed will principally turn on the degree to which the sizes of difference between the forces causing outcome rates of advantaged and disadvantaged to differ varies in the two situations and the degree to which the prevalence of the outcome varies in the two situations.

that fall below or above various points on a continuum of a quantifiable factor associated with experiencing some favorable outcome and the corresponding adverse outcome or simply showing the proportions of different groups that fall into various categories reflecting quantification of a factor.

For example, income data on which Table 2 is based show how lowering an income requirement to receive some favorable outcome will tend to increase relative difference in rates of failure to meet the requirement while reducing relative differences in rates of meeting the requirement.<sup>28</sup> Credit score data show the same thing with respect to a credit score requirement.<sup>29</sup> National Health and Nutrition Survey data show how generally reducing systolic blood pressure tends to increase relative racial differences in hypertension while reducing relative racial differences in rates of avoiding hypertension or that generally improving folate levels will tend to increase relative racial differences in low folate while reducing relative racial differences in rates of adequate folate.<sup>30</sup> Even data on rates at which groups fall into categories of health literacy show that the lower the category, the greater the relative difference in failing to achieve it and the smaller the relative difference in achieving it.<sup>31</sup> Data on rates at which groups meet certain proficiency levels will show the same thing.<sup>32</sup> And life tables show that the lower the age (and hence the lower are the rates of failing to survive to it) the larger tends to be the relative difference in failing to survive to the age while the smaller tends to be the relative difference in surviving to the age.<sup>33</sup>

As mortality declines relative difference in mortality tend to increase while relative differences in survival tend to decrease. As healthcare improves relative differences in receipt of appropriate care tend to decrease while relative differences in non-receipt of such care tend to increase. Generally decreasing any adverse outcomes like mortgage rejection, mortgage foreclosure, suspension or expulsion from school, dropping out of school, arrests, low birth weight tends to increase relative differences in rates of experiencing the outcome while reducing relative differences in avoiding it.

Similarly, relative racial and other relative differences in adverse outcomes tend to be larger, while relative racial differences in the corresponding favorable outcomes tend to be smaller, among advantaged populations/subpopulations/settings with comparatively low adverse

---

<sup>28</sup> See Table of the 2006 Chance [editorial](#).

<sup>29</sup> See the [Credit Score Illustrations](#) subpage of the [Scanlan's Rule](#) page of [jpscanlan.com](#).

<sup>30</sup> See the [NHANES Illustrations](#) subpage of the Scanlan's Rule page

<sup>31</sup> See the [Numeracy Illustration](#) subpage of the Scanlan's Rule page

<sup>32</sup> Rate of achieving a certain level of some sort of proficiency (including anything below it) should not be confused with rates of falling into a certain proficiency category. Rates at which certain groups fall into any intermediate category of an outcome – like rates at which different groups of students receive the grade of C – can never be effectively analyzed. See the [Intermediate Outcomes](#) subpage of the [Scanlan's Rule](#) page of [jpscanlan.com](#).

<sup>33</sup> See the [Life Tables Illustrations](#) subpage of the Scanlan's Rule page and the

outcome rates (*e.g.*, the young, persons receiving better healthcare like the insured and the military, the well-educated, persons with higher income, generally higher socioeconomic groups, states like Minnesota and Massachusetts) than among populations/subpopulations/settings with comparatively high adverse outcome rates.

For example, relative racial differences in infant mortality and low birth weight tend to be larger, while relative racial difference in infant survival and normal birth weight tend to be smaller among the well-educated and other low risk groups (where such outcomes are comparatively uncommon) than among high risk groups.<sup>34</sup> Relative racial differences in rates of mortgage rejection, mortgage foreclosure, and poor self-rated health tend to be greater, while relative racial differences in rates for the corresponding favorable outcomes tend to be smaller, among higher-income than among lower-income groups.<sup>35</sup> Relative racial differences in adverse discipline outcomes tend to be larger, while relative differences in the avoidance of such outcomes tend to be smaller, in settings where suspensions are less common (*e.g.*, schools without zero tolerance policies, suburban schools, states like Massachusetts with generally low suspension rates, pre-school) than in the corresponding settings where suspensions are more common.<sup>36</sup> Similarly, relative racial differences in adverse discipline outcomes tend to be larger, while relative racial differences in avoiding those outcomes tend to be smaller, among girls than among boys; correspondingly, relative gender differences in adverse discipline outcomes tend to be larger, while relative gender differences in avoiding those outcomes tend to be smaller, among whites than blacks.<sup>37</sup>

With only minor exception, the above patterns are utterly unknown in the law and the social and medical sciences. In fact, particularly with regard to infant and cancer outcomes, researchers will often refer to survival and mortality interchangeably, often stating they are examining relative differences in the former while in fact examining relative differences in the latter. They do so without recognizing the possibility, much less the likelihood, that the two relative differences will provide opposite results as to whether some demographic disparity is increasing or decreasing over time or is larger in one setting than another or with regard to one condition than another. See the [Mortality and Survival](#) page of [jpscanlan.com](#).

And a large proportion of persons analyzing demographic differences expects the exact opposite of the above patterns. That is, a large proportion of such persons expects that general reductions in an adverse outcome should reduce relative demographic differences in rates of experiencing the outcome and that relative racial/ethnic difference in adverse outcomes will tend

---

<sup>34</sup> See "[Race and Mortality](#)" and the unpublished paper "[The Relationship Between Declining Mortality and Increasing Racial and Socioeconomic Disparities in Mortality](#)."

<sup>35</sup> See the [Disparities – High Income](#) and [Foreclosure Disparities](#) subpage of the [Lending Disparities](#) page of [jpscanlan.com](#) and Figure 5 (slide 66) of the UMMS seminar.

<sup>36</sup> See the [Suburban Disparities](#), [DOE Equity Report](#), and the [Preschool Disparities](#) subpages of the [Discipline Disparities](#) page; letter to [Boston Lawyers' Committee for Civil Rights and Economic Justice](#) (Nov. 12, 2015); Table 8 of "Race and Mortality Revisited" (at 342).

<sup>37</sup> See Tables 1 and 2 of the [Discipline Disparities](#) page.

to be smaller within advantaged than disadvantaged subpopulations or settings. See especially [“Race and Mortality Revisited,”](#) *Society* (July/Aug. 2014) and [“Race and Mortality,”](#) *Society* (Jan./Feb. 2000).

That, of course, is the precise issue addressed in the shorter items discussed in the Introduction, which involve the longstanding belief of the government that reducing adverse lending, school discipline, criminal justice, and employment outcomes will tend to reduce relative racial/ethnic and other demographic differences in rates of experiencing those outcomes. But the belief persists because it is shared by a larger proportion of the scientific community, possibly close to all it. And the belief exists within the scientific community with regard to every type of adverse outcome.

There are two corollaries to the above-described pattern by which relative differences tend to be affected by the prevalence of an outcome.<sup>38</sup> One corollary is best described in the following terms with reference to a universe comprised of two groups: reducing the frequency of an outcome tends cause the group more susceptible to an outcome to make up a larger proportion of persons experiencing the outcome, and a larger proportion of persons failing to experience, than it previously did. For example, reducing poverty will tend to cause poorer groups to make up both a larger proportion of the poor and of the non-poor, as shown, for example, in Table 1 of the 2006 *Chance* editorial (which underlies Table 2 below). Reducing public school suspensions and adverse criminal justice outcomes will tend to increase the proportion racial minorities and other disadvantaged groups make up of persons experiencing those outcome (as well as increase the proportions such groups make up of persons failing to experience such outcomes).<sup>39</sup>

The appraisal of demographic disparities based on differences between the proportion a group comprises of persons experiencing an outcome and the proportion it comprises of persons actually experiencing the outcome is especially common in recent discussions of school discipline and criminal justice outcomes. Almost universally discussions of such differences reflect an expectation that reducing the frequency of adverse outcomes will tend to decrease the proportion disadvantaged groups make up of persons experiencing those outcomes. In addition to the 2016 items in *The Hill* and the *Federalist Society Blog* mentioned in the Introduction, see the letters to [Oklahoma City School District](#) (Sept. 20, 2016), [Antioch Unified School District](#)

---

<sup>38</sup> I sometimes refer to these corollaries as manifestations of the pattern rather than corollaries. I have also varied the order in which I present them. Compare the October 2015 ASA letter (at 9-10) with the UMMS seminar (slides 72-80).

<sup>39</sup> The pattern whereby reducing the frequency of an adverse outcome tends to increase the proportions disadvantaged groups make up of persons experiencing it and failing to experience it could just as well be termed the cause of the pattern of relative differences. For it is the fact that the disadvantaged group tends to comprise up a larger proportion of persons below and above each increasingly lower point on a continuum of a quantifiable factor associated with the likelihood of experiencing the favorable outcome that underlies the described pattern of relative differences. See Table 1 of the 2006 *Chance* [editorial](#). My earliest treatments of this subject principally addressed perceptions about disproportionate representations of certain groups among persons experiencing an adverse outcome. See [“The ‘Feminization of Poverty’ is Misunderstood,”](#) *Plain Dealer* (Nov 11, 1987); [“An Issue of Numbers,”](#) *National Law Journal* (Mar. 5, 1990) and [“The Perils of Provocative Statistics,”](#) *Public Interest* (Winter 1991).

(Sept. 9, 2016), [American Statistical Association II](#) (July 25, 2016), [City of Boulder, Colorado](#) (Mar. 5, 2016), [Houston Independent School District](#) (Jan. 5, 2016), [McKinney, Texas Independent School District](#) (Aug. 31, 2015), [Department of Health and Human Services and Department of Education](#) (Aug. 24, 2015), [City of Minneapolis, Minnesota](#) (June 8, 2015), [Texas Appleseed](#) (Apr. 7, 2015), and [United States Department of Justice and City of Ferguson, Missouri](#) (Mar. 9, 2015). As explained above, that expectation is the exact opposite of reality. The fact that, even apart from the implications of the patterns described here, it is impossible to soundly analyze demographic differences on the basis of a comparison of the proportion a group comprises of persons potentially experiencing an outcome and the proportion it comprises of persons actually experiencing the outcome, is discussed in Section C *infra*.

A second corollary to the described pattern of relative differences is that as an outcome changes in prevalence, including when a factor or intervention causes an outcome rate to change, the group with the lower baseline rate for the outcome will tend to experience a larger proportionate change in the outcome than the other group while the other group will tend to experience a larger proportionate change in the opposite outcome. As discussed previously, the fact that relative difference in mortgage rejection rates (or poor self-rated health) tend to be larger, while relative racial differences in mortgage approval rates (or good self-rated health) tend to be smaller, among higher-income groups than lower-income groups may also be regarded as reflecting the fact that having high income tends to causes a larger proportionate reduction in mortgage rejection rates (or rates of poor self-rated health) among whites than blacks while causing a larger proportionate increase in mortgage approval rates (or rates of good self-rated health) among blacks than whites. Every analysis of subgroup effects based on the comparative size of a relative effect (which is to say almost every analysis of subgroup effects/interaction) has been fundamentally unsound as a result of the failure to understand this pattern. See "[Race and Mortality Revisited](#)" at 339-341. The fact that, irrespective of the patterns described here, it is illogical to identify a subgroup effect on the basis of differing proportionate effects on different baseline rates is discussed in Section D *infra*.

As with the described pattern of relative differences, because other factors are also influencing observed patterns, the two corollaries will not always be observed. But in all situations where the described pattern of relative differences is observed, the corollaries will necessarily be observed as well.

\*\*\*

The pattern whereby the rarer an outcome the greater tends to be the relative difference in experiencing it and smaller tends to be the relative difference in avoiding it, as well as its two corollaries, can be easily illustrated with normally distributed test score data. Table 1 below, which is a version of Table 1 of "[Race and Mortality Revisited](#)" (at 329) and which reflects the same hypothetical employed in the shorter articles listed in the Introduction, is based on a situation where the means of normal test score distributions of an advantaged group (AG) and a disadvantaged group (DG) differ by half a standard deviation and both distributions have the same standard deviation. In addition to showing the pass and fail rates of each group, the table shows the ratio of AG's pass rate to DG's pass rate and the ratio of DG's fail rate to AG's fail

rate at each cutoff (the third and fourth last columns).<sup>40</sup> Based on a situation where AG and DG each make up half of the test takers, the final two columns show the proportion DG makes up of persons who pass and persons who fail at each cutoff.

**Table 1. Illustration of effects on relative differences in pass and fail rates of lowering a cutoff from a point where 80% of AG passes to a point where 95% of AG passes, with proportions DG comprises of persons who pass and of persons who fail (when mean scores differ by approximately half a standard deviation and DG comprises 50% of test takers)**

Cutoff	AG Pass	DG Pass	AG Fail	DG Fail	AG/DG Pass Ratio	DG/AG Fail Ratio	DG Prop of Pass	DG Prop of Fail
High	80%	63%	20%	37%	1.27	1.85	44%	65%
Low	95%	87%	5%	13%	1.09	2.60	48%	72%

According to the specifications underlying the table, at the cutoff where 80% of AG passes the test, approximately 63% of DG would pass the test (with corresponding failure rates of 20% for AG and 37% for DG). The ratio of AG’s pass rate to DG’s pass rate would be 1.27 while the ratio of DG’s fail rate to AG’s fail rate would be 1.85.

When the cutoff is lowered to the point where the pass rate for AG is 95%, the pass rate for DG would be approximately 87% (with corresponding failure rates of 5% for AG and 13% for DG). The ratio of AG’s pass rate to DG’s pass rate would thus decrease to 1.09 (from 1.27), while the ratio of DG’s fail rate to AG’s fail rate would increase to 2.60 (from 1.85). That is, the relative difference in the outcome that is reduced in frequency (test failure) increases (from 85% to 160%), while the relative difference in the increasing outcome (test passage) declines (from 27% to 9%).<sup>41</sup>

The final two columns also illustrate the first corollary to the described pattern by which relative differences tend to be affected by the prevalence of an outcome. Lowering the cutoff and reducing the frequency of test failure caused an increase in the proportion DG makes up of those who pass the test (from 48% to 52%) and the proportion DG makes up of persons who fail the test (from 65% to 72%). Because the proportion DG makes up of persons taking the test is

<sup>40</sup> While I commonly refer to patterns of relative differences in these comments, the table actually presents rate ratios. 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. One should be careful not to mistakenly refer to the rate ratio as the relative difference. But the distinction between the two terms is not pertinent to the discussion here of patterns by which the two relative differences tend to be affected by the prevalence of an outcome. In recent years I commonly present the rate ratios for both outcomes with the larger figure in the numerator, in which case, as to both outcomes, the larger the rate ratio, the larger the relative difference. In the 1994 *Chance* [article](#) and the 2006 *Chance* [editorial](#) I used the disadvantaged group’s rate as the numerator in both ratios (which is the approach of the “four-fifths” or “80 percent” rule for identifying disparate impact under the Uniform Guideline for Employee Selection Procedures). Choice of numerator in the ratio, however, has no bearing on the patterns by which the two relative differences tend to be affected by the prevalence of the outcome.

<sup>41</sup> I commonly speak of the decreasing outcome first because failures of understanding typically involve that outcome. But, partly out of convention and partly to be consistent with certain earlier illustrations, Table 1 presents the increasing outcome first and discusses that outcome first save in the sentence to which this note is attached.

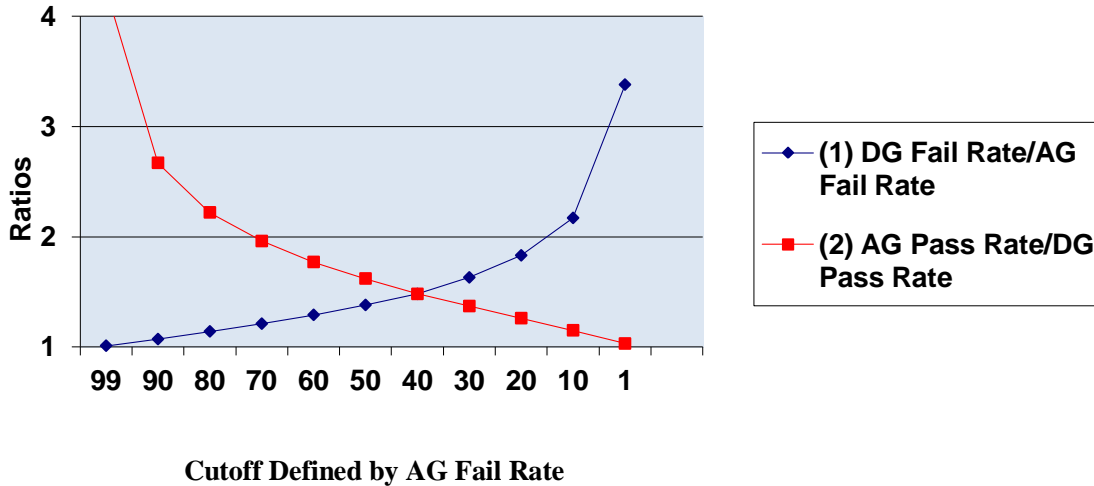
unaffected by the cutoff, lowering the cutoff would increase all measures of differences between the proportion DG makes up of test takers and the proportion it makes up of persons who fail (while reducing all measures of difference between the proportion DG makes up of test takers and the proportion it makes up of persons who pass).

The table also illustrates the second corollary to the described pattern by which relative differences tend to be affected by the prevalence of an outcome. Lowering the cutoff caused a larger proportionate decline in failure rates for AG (the group with the lower baseline failure rate) while causing a larger proportionate increase in the pass rate for DG (the group with the lower baseline pass rate). That is, lowering the cutoff caused failure rates to decrease by 75% for AG but only 65% for DG, while causing pass rates to increase by 38% for DG but only 19% for AG. Raising the cutoff back to the original point would similarly show a larger proportionate increase in the failure rate for AG but a larger proportionate decrease in the pass rate for DG, thus causing the relative difference in failure rates to decrease and the relative difference in pass rates to increase.

Table 1 can also illustrate the pattern whereby relative differences in the adverse outcome tend to be larger, while relative differences in the corresponding favorable outcome tend to be smaller, in advantaged populations/subpopulations/settings than in disadvantaged populations/subpopulations/settings. The reader need simply regard the second row as reflecting the situation in the advantaged population/subpopulation/setting (where the adverse outcome is less common) and the first row as reflecting the situation in the disadvantaged population/subpopulation/setting (where the adverse outcome is more common). In terms of the point of the preceding paragraph, this means that being in the advantaged population/subpopulations/setting tends to reduce adverse outcome rates proportionately more for advantaged groups than disadvantaged groups while increasing favorable outcome rates proportionately more for disadvantaged groups than advantaged groups. To make the last point more concrete, having high income or high education tends to reduce adverse outcome rates proportionately more for whites than disadvantaged racial minorities, while increasing favorable outcome rates proportionately more for disadvantaged racial minorities than whites.

The pattern of relative differences in pass and fail rates shown in Table 1 exists across the full range of test scores. Figure 1, which employs the same specifications as Table 1, shows the effects on the two relative differences of lowering a cutoff from a point where almost everyone fails to a point where almost every passes. The relative difference in the decreasing outcome (test failure) consistently increases, while the relative difference in the increasing outcome (test passage) consistently decreases.

**Figure 1. Ratios of (1) DG Fail Rate to AG Fail Rate, (2) AG Pass Rate to DG Pass Rate, at Cutoffs Defined by AG Failure Rates**



The two corollaries to the above pattern will similarly be observed across the full range of prevalence of the outcome.

Table 2, which is an abbreviated version of Table 2 or "Race and Mortality Revisited" and Table 1 of the 2006 *Chance* editorial, shows proportions of the white and black populations living on income below and above the poverty line and 75% of the poverty line, along with the ratios of the black to white ratios of falling below each point and the white to black ratios of falling above the point. The table also includes the percentage point difference between rates, though I will make only limited reference to it in this section.

**Table 2. Rates at which white and blacks fall above and below 125%, 100%, and 75% of the poverty line, with measures of differences between rates (2004)**

Percent of Poverty Line	Percent of Whites Below	Percent of Blacks Below	Percent of Whites Above	Percent of Blacks Above	B/W Below Ratio	W/B Above Ratio	Percentage Point Diff
100%	10.8%	24.7%	89.2%	75.3%	2.29	1.18	14
75%	7.2%	17.8%	92.8%	82.2%	2.48	1.13	11

Movement from the first to the second row shows the implications of reducing poverty such as to enable everyone living on incomes above 75% of the poverty line to escape poverty. The black-white poverty ratio would increase from 2.29 to 2.48 (i.e., the relative difference would increase from 129% to 148%), while the ratio of the white rate of avoiding poverty to the black rates of avoiding poverty would decline from 1.18 to 1.13 (i.e., relative difference would decrease from 18% to 13%). And the absolute difference would decline from 14 to 11 percentage points.



As shown in Table 2 of "Race and Mortality Revisited" and Table 1 of 2006 *Chance* editorial, one would observe the opposite patterns of changes in measures if poverty were to increase such as to pull into poverty all persons living on incomes below 125% of the poverty line. Such pattern may also be observed in Table 2 above if one simply envisions that after the initial reduction in poverty simulated by the movement from the first to the second row of the table, poverty rose to its prior level.

The table does not present the proportion blacks make up of the combined black and white population below and above each point. Such information may be found in Table 1 of the 2006 *Chance* editorial. But, as explained with regard to the test score illustration in Table 1 above, given the data in Table 2 above, the proportions blacks comprise of the combined black and white populations falling both below and above 75% of the poverty line would necessarily be greater than the proportions falling below and above the poverty line itself. As discussed in note 39, the fact that such is the case may be deemed the reason for the patterns of relative differences shown in the table.

In "Race and Mortality Revisited" I discuss with regard to its Table 2 whether there could be justification for devoting any resources to exploring why, during a period when there occurred a general increase or decrease in poverty, the relative difference between black and white poverty rates or rates of avoiding poverty changed – say, for example, to attempt to determine whether a particular administration's civil rights enforcement policies may have played some role – without consideration of the patterns in Table 2. The obvious answer is that there could be no justification.

The same point holds with regard to any exploration of reasons observed patterns of differences in the proportionate changes in the black and white poverty rates or rates of avoiding poverty, as might commonly be done in discussions that one group or another was disproportionately affected by the general change in poverty. The point also holds with regard to exploration of reasons for changes in the proportion blacks comprise of the poor or of the non-poor. And it would hold as well with regard to explorations of reasons for observed pattern of changes in absolute differences between poverty rates.

With regard to the potential focus on differences in the relative (or absolute) change in each group's rate in the case of a decline (or increase) in poverty or anything else, the following should be borne in mind. Given the manner in which demographic differences are commonly analyzed, in the case of the reduction in poverty simulated by the movement from the first to the second row of Table 2, it would not be surprising (a) for researchers exploring the cause of the increase in the relative difference in poverty rates to attribute the increase to the fact that the white poverty rate fell proportionately more than the black poverty rate or (b) for researchers exploring the cause of the decrease in the absolute difference to attribute the decrease to the fact that the black poverty rate fell by a greater absolute amount than the white rate. That in fact is what is done in the discussion of increasing relative socioeconomic differences, but declining absolute socioeconomic differences, in mortality in the BMJ article<sup>42</sup> that is the subject of my [Comment on Mackenbach BMJ 2016](#) (2016).

---

<sup>42</sup> Mackenbach JP, Kulhánová I, Artnik B, et al. Changes in mortality inequalities over two decades: register based study of European countries. *BMJ* 2016;353:i1732.

Some observers might regard such information as providing insight into causes of increasing or decreasing disparities. But as with discussion of the changes in the size of the disparities themselves, such discussion is providing nothing about whether the observed pattern is anything other than the consequence of the change in the prevalence of the outcome. That is, such discussion merely describes a mathematical relationship; it provides no information about whether the differences in the circumstances of the groups have actually changed or about the role of policies in effecting changes in the difference in the circumstances of the groups.

As suggested above, however, research examining patterns of changes in demographic differences in poverty or any other matter has never taken patterns like those in reflected in Table 2 into account. Thus, it has not attempted (that is, attempted in a sound manner) to identify things that it might actually be useful for policymakers to know.

\*\*\*

In the main Introduction to these comments and the introduction to this part, I mentioned that the National Center for Health Statistics (NCHS) came to understand more than a decade ago that determinations of whether health and healthcare disparities were increasing would commonly turn on whether one examined relative differences in a favorable outcome or relative differences in the corresponding adverse outcome. This occurred as a result of my bringing to the attention of NCHS statisticians in 2002 “[Race and Mortality](#),” *Society* (Jan./Feb. 2000) and “[Divining Difference](#),” *Chance* (Fall 1994). This was not, however, the agency’s first recognition of the pattern. As discussed at the end of “Race and Mortality” the director of the agency had come to recognize the pattern several years earlier, which had occurred as a result of my contacts to the directors of the Race and Health Initiative.<sup>43</sup> But as of 2002 the information had not found its way to the NCHS statisticians who developed policy on the measurement of health disparities.

In any case, the communications directly to NCHS statisticians led to their publishing four official or unofficial papers between 2004 and 2009 discussing that determination of directions of change in health and healthcare disparities would commonly turn on which relative difference one examined. The most important of these was a 2005 monograph titled “[Methodological Issues in Measuring Health Disparities](#)” (authored by NCHS personnel and other experienced health disparities researchers).<sup>44</sup> Each of the documents simply concluded that

---

<sup>43</sup> See the January 20, 1999 [letter](#) from NCHS Director Edward J. Sondik. See also my January 25, 1999 [letter](#) to Director Sondik explaining that the absolute difference between rates did not provide a solution to the problem of the contradictions in interpretations provided by the two relative differences because the absolute difference tended also to be affected by the prevalence of an outcome.

<sup>44</sup> Keppel K, Pamuk E, Lynch J, et al. Methodological issues in measuring health disparities. *Vital Health Stat.* 2005;2(141):116. Other papers include: (a) Keppel KG, Percy JN, Klein RJ. 2004. [Measuring progress in Healthy People 2010](#). Healthy People statistical notes. No. 25. Hyattsville, Md.: National Center for Health Statistics; (b) Keppel KG, Percy JN. 2005. Measuring relative disparities in terms of adverse events. *J Public Health Manag Pract*, 11(6):479–483; and (c) Keppel KG, Percy JN. 2009, Healthy People 2010: Measuring disparities in health.

in order to promote consistency, for purposes of measuring health disparities reduction goals in Healthy People 2010, both health and healthcare disparities would be measured in terms of relative differences in adverse outcomes (meaning, in the case of healthcare, relative differences in non-receipt of care).

I have in many places criticized the NCHS manner of dealing with this issues (both as to health and healthcare) for the failure to recognize that the purpose of health and healthcare disparities research is to understand processes and inform policies, which cannot be achieved by arbitrarily choosing a measure that tends to indicate that disparities are increasing over one that tends to indicate that they are decreasing (or vice-versa). As indicated above, that purpose can only be achieved by taking into account the ways the measures employed in such research tend to be affected by the prevalence of an outcome.

Some implications of the NCHS approach are discussed in "[Race and Mortality Revisited](#)" and the FCSM [paper](#), especially with regard to (a) a 2008 study where, ignoring or unaware of NCHS guidance, the authors measured disparities in terms of relative difference in receipt of immunization and found a dramatic increase in immunization rates to be associated with dramatically reduced racial/ethnic disparities (while NCHS, relying on relative differences in non-receipt of immunization, would have reached opposite conclusions) and (b) a 2009 study where, specifically relying on NCHS guidance, the authors measured disparities in terms of relative differences in non-receipt of mammography and found dramatic increases in mammography rates to be associated with dramatic increases in area socioeconomic disparities (while those relying on relative differences in receipt of mammography would have reached opposite conclusions). See discussion of Table 4 in "'Race and Mortality Revisited'" and Tables 3 and 7 in the FCSM paper. The two papers discuss a number of similar situations where analyses of changes in healthcare disparities would reach opposite results depending on whether researchers followed or ignored (usually meaning were unaware of) NCHS guidance, as well as a situation where NCHS's views as to the largest and smallest disparities in skilled birth attendance in certain low and middle income countries would be the opposite of the views of the World Health Organization. See Table 4 of the FCSM paper.

As discussed in the 2016 JPHMP [commentary](#), however, the NCHS has recently reversed its position regarding the measurement of healthcare disparities and now measures them in terms of relative differences in favorable outcomes. Thus, NCHS would now agree with the authors of the 2008 immunization study who ignored the earlier guidance and disagree with the authors of the 2009 mammography study who specifically followed NCHS guidance. Reversals of interpretation would apply to all situations discussed in "Race and Mortality Revisited" and the FCSM paper regarding healthcare disparities interpretation issues as to which the position taken in the 2005 NCHS monograph is pertinent.

Another illustration of the implications of the NCHS reversal of position with regard to researchers who relied on the guidance may be found at pages 5-6 of the March 8, 2016 letter to [Stanford Center on Poverty and Inequality](#) regarding a *Poverty and Inequality Report* titled

---

*Chance*, 22(1):6-9. See also Keppel K.G., Pearcy J.N. 2006. [Response to Scanlan concerning: measuring health disparities in terms of adverse events](#). *J Public Health Manag Pract*, 12(3):295.

“[State of the Union](#)” that the Center issued in early 2016. The letter discusses the way that authors of a portion of the report relied on NCHS guidance to measure demographic differences in insurance coverage in terms of relative difference in uninsurance and drew inferences about underlying processes based on the comparative size of those differences in different settings. Prior to issuance of the report, NCHS had already reversed its position such that (though the data were unavailable to confirm the pattern) observers relying on the revised guidance would commonly draw very different inferences about processes from those drawn by authors of the Stanford Center’s report.

The most notable consequence of the NCHS reversal, however, is that it constitutes a repudiation of more than a decade of yearly National Healthcare Disparities Reports issued by the Agency for Healthcare Research and Quality (AHRQ). Following NCHS guidance, the report had relied on the guidance to measure healthcare disparities in terms of relative differences in non-receipt of care. This issue and whether AHRQ is yet aware of the reversal or its implications is discussed further in Section B.

While a Google search indicates that the 2005 NCHS monograph has been cited 180 times, to my knowledge, none of the papers citing it or citing other items in which NCHS has recommended measuring healthcare disparities in terms of relative differences in adverse outcomes (excluding situations where I have cited them) has indicated an awareness that reliance on relative differences in adverse outcomes will commonly result in opposite conclusions from reliance on relative differences in the corresponding favorable outcomes.<sup>45</sup> Other federal agencies involved with health and healthcare disparities research have yet to show an awareness that it is even possible for the two relative differences to yield opposite conclusions about patterns of changes in health and healthcare disparities. In consequence of the retirement of the principal NCHS author of the 2005 monograph, it is questionable whether more than a few persons at NCHS are aware that NCHS statisticians ever reached the conclusion they did in the monograph and the three other papers.

Most important, neither NCHS nor any other agency conducting research into health and healthcare disparities, providing guidance on such research, or funding such research has considered the implications of the prevalence of an outcome on measures employed in such research. The same holds for the private sector.

\*\*\*

As discussed in the August 24, 2015 [letter](#) to the Department of Health and Human Services (HHS) and Department of Education (DOE), the Secretary of HHS has a view as to the effects or reducing adverse outcomes on relative differences in rates of experiencing those outcomes that is both the opposite of the view NCHS reached a decade ago and the opposite of

---

<sup>45</sup> See pages 26-27 of the FCSM paper regarding Healthy People 2010 guidance on measurement of healthcare disparities indicating that healthcare disparities will be discussed in terms of favorable outcome even though they are being measured in terms of relative differences in adverse outcomes and the fact that few readers would infer from the language that the relative difference in the adverse outcome may well be yielding an opposite conclusion from the relative difference in the favorable outcome. See also the slide 32 of the [presentation](#) associated with the FCSM paper.

reality. Specifically, the Secretary of HHS, like the Secretary of Education, believes that reducing public school suspensions will tend to reduce demographic differences in discipline rates and the proportions disadvantaged groups make up of persons experiencing those outcomes.

The reasons that such actions will tend to have the opposite effect have been explained above. In addition, in point of fact, recent reductions in discipline rates have consistently been accompanied by increased relative differences in discipline rates (and thus the proportion more susceptible groups make up of those disciplined) notwithstanding that teachers and administrators are no doubt taking various actions that would be expected to reduce all measures of differences between outcome rates. See the subpages to the subpages of the [Discipline Disparities](#) page of [jpscanlan.com](#) discussing the situation with regard to the jurisdictions indicated in the title of the subpage: [California Disparities](#), [Colorado Disparities](#), [Connecticut Disparities](#), [Maryland Disparities](#), [Minnesota Disparities](#), [Oregon Disparities](#), [Beaverton, OR Disparities](#), [Denver Disparities](#), [Henrico County, VA Disparities](#), [Los Angeles SWPBS](#), [Minneapolis Disparities](#), [Montgomery County, MD Disparities](#), [Portland, OR Disparities](#), [St. Paul Disparities](#).<sup>46</sup> See also the [DOE Equity Report](#) subpage regarding a Department of Education report showing that relative racial differences in expulsions are larger in school districts without zero tolerance policies than school districts with such policies; the [Suburban Disparities](#) subpage regarding the greater relative racial differences in suspensions in suburban than central city schools; "Race and Mortality Revisited" (at 342) regarding the larger relative differences in suspension rates in pre-school than in K-12; and the Boston Lawyers' Committee letter mentioned in note 6 regarding larger relative differences in suspensions in Massachusetts (which has generally low suspension rates) than nationally.

Some of the problems faced by particular entities as a result of the mistaken belief of the government or others that reducing the prevalence of adverse discipline outcomes will tend to reduce relative differences in rates of experiencing the outcomes, and the proportions disadvantaged groups make up of persons experiencing the outcomes, are discussed in the [Oakland Agreement](#) page and the letters to school districts listed on pages 18-19. See especially the letters to [Oklahoma City School District](#) (Sept. 20, 2016) regarding a recent agreement with the Department of Justice, the [Antioch Unified School District](#) (Sept. 9, 2016) regarding a recent suit brought against that district by a public interest group, and [McKinney, Texas Independent School District](#) (Aug. 31, 2015) regarding a public interest group's analyses of racial/ethnic differences in the district.

Similar problems faced by particular entities as a result of the government's beliefs about the effect of generally reducing criminal justice outcomes on measures of racial disparity may be found in the letters to cities listed on pages 18-19. The particular problems facing Ferguson, Missouri are discussed in the letter to [United States Department of Justice and City of Ferguson](#),

---

<sup>46</sup> Reportage of situations where general reductions in discipline rates have been accompanied by reduced racial differences in discipline have generally pertained to studies that measured disparities in terms of absolute differences between rates. As discussed in the letter to the Antioch Unified School District, absolute differences between rates tend to decline when outcomes in the rate ranges commonly observed for adverse school discipline outcomes generally decline.

[Missouri](#) (Mar. 9, 2015) and in the “[Things DoJ doesn’t know about racial disparities in Ferguson](#),” *The Hill* (Feb. 22, 2016). See also my [Submission re Ferguson Consent Decree](#) (Apr. 11, 2016) regarding specific issues in the consent decree resolving the Department of Justice’s suit against the city. The recent “[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) gives an indication of similar problems facing Baltimore, Maryland.

\*\*\*

A few concluding points are in order regarding the two relative differences. Congress has no better an understanding of these issues than the federal law enforcement agencies or the federal agencies involved with health and healthcare disparities research. In fact, as discussed in "Race and Mortality Revisited" (at 342) and the [Disabilities – Public Law 104-446](#) subpage of the Discipline Disparities page, the Individuals with Disabilities Education Act mandates that recipients of federal assistance with “significant discrepancies” in long-term suspensions of students with and without disabilities, which are commonly measured in terms of relative differences in suspension rates, must consider approaches to discipline of the type that generally reduce suspension rates. The implementation of those approaches thus will tend to increase the discrepancies according to the standard approach to measuring them. See the [Keep Kids in School Act](#) subpage of the Discipline Disparities regarding proposed legislation contemplating that, in accordance with beliefs of the Department of Justice, Education, and Health and Human Services, generally reducing discipline rates will reduce relative racial differences in discipline rates.

The Civil Rights Act of 1991 provides that employers must justify practices having a disparate impact on protected groups and implement less discriminatory alternatives to practices that can be justified. In enacting the statute Congress had no understanding that relaxing standards or otherwise reducing adverse employment outcomes, and increasing the corresponding favorable outcomes, would tend to increase relative differences in the adverse outcomes while reducing relative differences in the favorable outcomes. Congress continues to have no understanding that it is even possible for the two relative differences to change in the opposite outcome as the prevalence of an outcome changes. Meanwhile, it remains possible for courts either to measure disparate impacts in terms of relative differences in favorable outcome or relative differences in adverse outcomes. No court has yet recognized that it is possible for the two relative differences to change in opposite directions as the prevalence of an outcome changes. See the letter to [Federal Judicial Center](#) (July 7, 2016). See the [Less Discriminatory Alternative - Substantive](#) subpage of the [Disparate Impact](#) page. See generally the main Disparate Impact page and each of its subpages regarding the lack of understanding of the relationship between the frequency of an outcome and measures of differences between outcome rates among all entities having a role in the enforcement of laws against employment discrimination.

Federal laws impose a variety of obligations to monitor demographic differences in certain outcomes. The obligation to monitor demographic differences in assignment to special education programs imposes what would seem to be an very large administrative burden on

states. Yet, neither the Congress that imposes the obligations, the Department of Education that promulgates regulations implementing or expanding those obligations, the Government Accountability Office that attempts to monitor the actions of the Department of Education, nor the divisions of the state governments that attempt to fulfill those obligations are yet aware that actions that generally reduce assignment to special education tend to increase relative racial differences in assignment rates.

In sum, the matters discussed above only suggest the scope of activities of government and private actors where failure to understand the ways relative differences in outcome rates tend to be affected by the prevalence of an outcome undermine programs aimed at monitoring or ameliorating demographic differences.

## **B. Absolute Differences and the Odds Ratios**

Appraisals of differences in the circumstances of advantaged and disadvantaged groups reflected by their outcome rates in terms of absolute (percentage point) differences between rates and differences measured by odds ratios are unaffected by which outcome one examines. But in order for a measure to effectively quantify the strength of the forces causing outcome rates to differ, it must remain unchanged as there occurs a general change in the prevalence of an outcome akin to that effected by the lowering of a test cutoff. And, like the two relative differences, absolute differences and odds ratios also tend to be affected by the frequency of an outcome, though in a more complicated way than the two relative differences.

Roughly, as uncommon outcomes (below 50% for both groups) become more common, absolute differences between rates tend to increase; as common outcomes (above 50% for both groups) become even more common, absolute differences tend to decrease. The frequency-driven direction of change of the absolute difference is harder to predict when the outcome is neither common nor uncommon or moves between categories of prevalence during a period examined. In the situations reflected in Tables 1 and 2, the rates of experiencing the favorable and corresponding adverse outcomes happen to be in ranges where movements from the first row to the second row would reduce absolute differences, while movement from the second row to the first row would increase absolute differences. The school discipline and criminal justice outcomes where focus is generally on relative differences in adverse outcomes or the proportions disadvantaged groups make up of persons experiencing those outcomes – both of which tend to increase as the outcome generally decreases – also happen to be in ranges where reductions in the adverse outcome will tend to reduce absolute differences between rates.

The absolute difference and both relative differences may all change in the same direction as the frequency of an outcome changes (in which case the difference measured by the odds ratio will also change in the same direction as the other measures). But in the common situation where all measures do not change in the same direction as the frequency of an outcome changes, the absolute difference will tend to change in the same direction as the smaller relative difference. Observers who rely on relative differences to appraise the difference in the circumstances of two groups reflected by their differing outcome rates tend

usually to rely on the larger of the two relative differences (as discussed above).<sup>47</sup> Thus, there exists a systematic tendency for observers relying on the absolute difference to reach opposite conclusions about directions of changes in the size of demographic differences, or about the comparative size of the differences in different settings, from observers relying on a relative difference.

Further, anytime an observer notes that the absolute difference has changed in a different direction from the relative difference the observer happens to be examining, the unmentioned relative difference will necessarily have changed in the opposite direction of the mentioned relative difference and the same direction as the absolute difference. Thus, observers maintaining that one must make a value judgment in choosing between the relative difference and the absolute difference with respect to the appraisal of a particular change in some demographic disparity over time have already made a choice (usually without thought) to rely on the relative difference that yields an opposite conclusion from the absolute difference rather than the relative difference that yields the same conclusion as the absolute difference.

As the frequency of an outcome changes, and all measures do not change in the same direction, the difference measured by the odds ratio tends to change in the opposite direction of the absolute difference and in the same direction as the larger relative difference.

These patterns are illustrated in Figures 2 and 3, which are based on the same specifications as Figure 1, and similarly show the implications of lowering a cutoff from a point where almost everyone fails to a point where almost everyone passes. Figure 2 presents the pattern for the absolute difference. Figure 3 presents the pattern for the ratio of the disadvantaged group's odds of failing the test to the advantaged group's odds of failing the test.<sup>48</sup> But in order to illustrate the relationship with the two relative differences, a line for the odds ratio is simply added to the lines for the two relative differences previously shown in Figure 1.

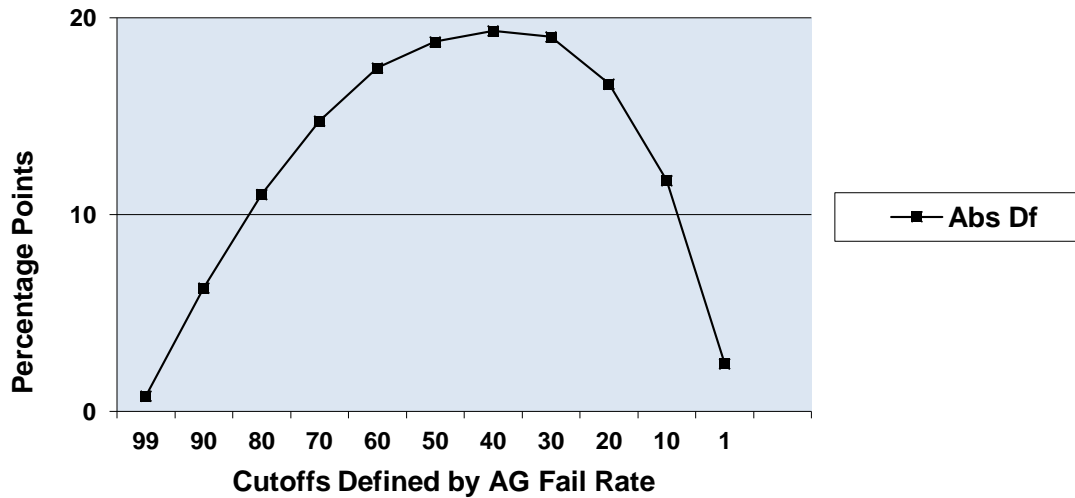
---

<sup>47</sup> The tendency to rely on the larger of the two relative differences sometimes causes researchers to rely both on relative differences in favorable outcomes and on relative differences in adverse outcomes in the same study. See the [Immunization Disparities](#) page of [jpscanlan.com](#) regarding a study that examined immunization disparities in terms of relative differences in the adverse outcome for receipt/non-receipt of any immunization and relative differences in the favorable outcome for receipt/non-receipt of full immunization and the [McKinsey Achievement Gap Study](#) subpage of the [Educational Disparities](#) page regarding a study where the authors relied on the relative difference in the adverse outcome for reaching/failing to reach the basic proficiency level but on the relative difference in the favorable outcome for reaching/failing to reach the advanced proficiency level. In both cases the authors relied on the larger of the two relative differences with respect to each of the subjects examined and without recognizing that general increases in favorable outcome would tend to increase relative differences in the adverse outcomes while reducing relative differences in the favorable outcomes. See also the letter to [New York City Center for Innovation through Data Intelligence](#) (June 6, 2016).

<sup>48</sup> There are four odds ratios depending on which outcome is examined and which group's odds is used as the numerator of the ratio. Two yield one value and two yield a value that is the reciprocal of the first value. The ratio of DG's failure odds to AG's failure odds is the same as the ratio of AG's pass odds to DG's pass odds. Thus, the odds ratios underlying Figure 3 are also the ratios of AG's odds of passing the test to DG's odds of passing the test.



**Figure 2. Absolute differences between rates of AG and DG pass (or fail) rates at various cutoff points defined by AG fail rate**



**Figure 3. Ratios of (1) DG fail rate to AG fail rate, (2) AG pass rate to DG pass rate, (3) DG failure odds to AG failure odds**

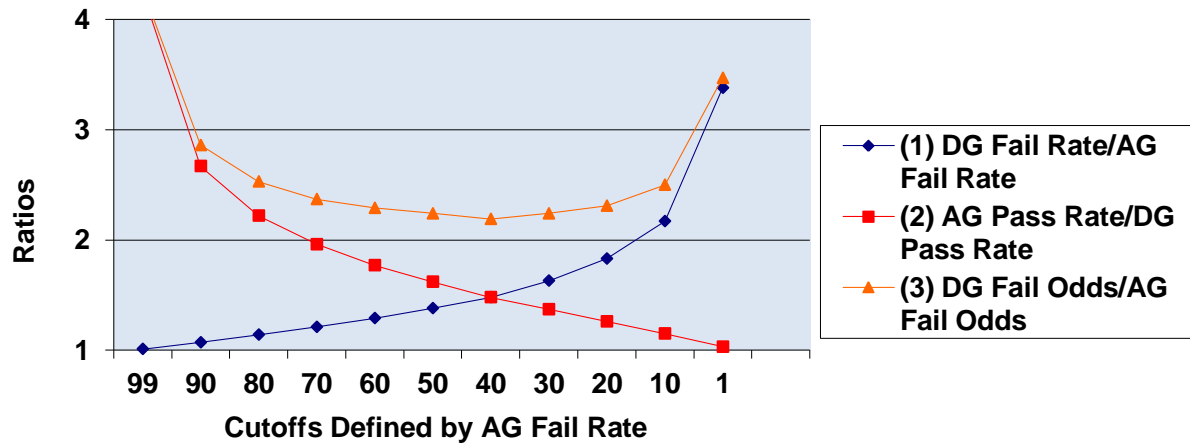


Table 3 below, which is based on the same specifications as Table 1 and Figures 1 to 3, presents favorable outcome rates at four prevalence levels (benchmarked on the advantaged group’s favorable outcome rate), along with rate ratios for the favorable and adverse outcomes as well as absolute differences and odds ratios. The parenthetical numbers indicate the ways one would rank the size of the disparity between the situation of the advantaged and disadvantaged groups according to each of the four measures. Those numbers show that rankings according to

relative differences in favorable outcomes are the opposite of rankings according to relative differences in adverse outcomes. They also show that rankings according to absolute differences and odds ratios, while different from the rankings according to either relative difference, are the opposite of each other.

**Table 3. Favorable outcome rates of advantaged group (AG) and disadvantaged group (DG) at four settings with different favorable outcome frequencies, with measures of difference**

	<b>AG Fav Rate</b>	<b>DG Fav Rate</b>	<b>AG/DG Fav Ratio</b>	<b>DG/AG Adv Ratio</b>	<b>Abs Df (Perc Pnts)</b>	<b>Odds Ratio</b>
A	20.0%	9.0%	2.22 (1)	1.14 (4)	11.0 (4)	2.53 (1)
B	40.0%	22.6%	1.77 (2)	1.29 (3)	17.4 (2)	2.28 (3)
C	70.0%	51.0%	1.37 (3)	1.63 (2)	19.0 (1)	2.24 (4)
D	80.0%	63.4%	1.26 (4)	1.83 (1)	16.6 (3)	2.31 (2)

A version of this table appears as Table 5 (at 335) of "Race and Mortality Revisited," where, as in many other places, it is used to refute the notion that choice of a measure in analyses of demographic differences in health or healthcare outcomes involves a value judgment. I will rest here on the treatment of that issue in "Race and Mortality Revisited" (at 335-336) and other places.

But I also note the following. In the case of Table 5 in "Race and Mortality Revisited" I cast the matter cast in terms of situations where one has to rank, according to the degree of discrimination or the likelihood of discrimination, four employers or one employer at four points in time. I cast the matter that way in order to cause the reader to recognize that there can be only one answer to the question of whether the forces causing the outcome rates to differ are larger in one situation than another and that, while the answer may be difficult to divine, value judgments have can have no role in the matter.

The same considerations apply when the purpose of examining patterns of differences in outcome rates is to inform policies aimed at mitigating differences in the circumstances of two groups. Further, anyone inclined to believe that the particular measure ought to be the focus in the appraisal of demographic differences irrespective of prevalence considerations ought to recognize that policymakers have a great interest in knowing whether policies will cause that measure to change more than, less then, or in a different direction from what would be likely to occur solely as a consequence of changes in the prevalence of an outcome.

But I include Table 3 here principally to illustrate certain points about the absolute difference in contexts where observers commonly rely on that measure. The premise of the table, as reflected by its specifications, is that there is no rational basis to maintain that the strength of the forces causing the favorable (or adverse) outcome rates of the advantaged and disadvantaged groups to differ varies among the rows.

Areas of research or commentary where observers commonly or increasingly rely on absolute differences between rates include appraisals of demographic differences in healthcare and academic proficiency.

In the case of healthcare, increases in rates of appropriate care will tend to increase absolute differences between rates of advantaged and disadvantaged groups for uncommon procedures/outcome, while reducing absolute differences for common procedures/outcomes. Such patterns are reflected by movement from Row A to Row B of Table 3 for the former procedures/outcomes and from Row C to Row D for the latter procedures/outcomes. When healthcare outcomes increase from being fairly uncommon to being very common (as in the case of certain vaccines and screening practices) the absolute differences will tend to increase for a time and then decrease.

Further, higher-performing hospitals (which tend to have generally higher appropriate care rates than lower-performing hospitals) will tend to show larger absolute differences between appropriate care rates of advantaged and disadvantaged groups than lower-performing hospitals for procedures/outcomes where rates are generally low, while showing smaller absolute differences for procedures/outcomes where rates are generally high. This pattern, too, is illustrated in Table 3 if Rows A and B are regarded as the lower- and higher-performing hospitals with respect to procedures/outcomes with generally low rates and Rows C and D are regarded as such hospitals with respect to procedures/outcomes with generally high rates.

Odds ratios would tend to show patterns that are the opposite of those just described for absolute differences. In situations where observers draw initial conclusions about such things as changes in the size of healthcare disparities over time on the basis of absolute differences between rates, and then employ logistic regression to adjust for the role of possible confounders, the odds ratio yielded by the adjustment will tend to show a pattern that is the opposite of that shown by the absolute difference.<sup>49</sup>

Observers relying on absolute differences to measure healthcare disparities, however, have yet to recognize the ways absolute differences tend to be affected by the frequency of an outcome or that other measures would tend to systematically yield opposite results (*i.e.*, the larger relative difference and the odds ratio) or consistent results (*i.e.*, the smaller relative difference). Commonly observers relying on absolute differences, like those relying on one of the two relative differences, show no awareness that choice of measure is of any consequence.

The confusion arising from the failure to understand patterns by which absolute differences between health and healthcare outcome rates of advantaged and disadvantaged groups tend to be affected by the prevalence an outcome is vast. "Race and Mortality Revisited" (at 337-339) discusses the way the failure to understand these patterns caused Massachusetts unwisely to include a health disparities element in its Medicaid pay-for-performance program and then to measure disparities in a way that, by generally favoring higher-performing hospital

---

<sup>49</sup> Though I give only limited attention to the odds ratio, I recognize that some might deem it a measure essentially unaffected by the prevalence of an outcome (even though it seems principally to be employed as a convenient substitute for a rate ratio). I discuss the potential for the odds ratio to be an effective measure in note 22 (at 14) of letter to the [Antioch Unified School District](#) (Sept. 9, 2016), but leave the utility of the measure, like the utility of the measure I discuss in "Race and Mortality Revisited," as a matter for the Commission to consider in its efforts to reform the analyses of demographic differences with regard to the way such analyses affect appraisals of policies and programs. I merely note here that, as with other measures, researchers should be required to address whether and how the differences measured by the odds ratio tends to be affected by the prevalence on an outcome in the case of the matter under study. The same holds for any regression approach to appraising demographic differences.

for reasons unrelated to actual differences in equity of care, will tend to increase healthcare disparities.

A useful illustration of the way prevalence issues are ignored may be found in a group of articles appearing in the August 18, 2015 issue of the *New England Journal of Medicine* (NEJM), along with a commentary thereon and a letter to the editors by three authors of the 2005 NCHS monograph. A study by Vaccarino *et al.*<sup>50</sup> relied on relative differences in favorable healthcare outcomes (though relative differences in adverse outcomes for health status issues) with regard to outcome rates that were not changing much in overall prevalence during the period examined; and, as commonly happens when overall prevalence does not change much, the study found little to remark on with respect to changes in disparities over time. A study by Jha *et al.*<sup>51</sup> relied on absolute differences between rate in examining disparities in rates of receiving certain fairly uncommon procedures that were generally increasing in overall prevalence; and, as commonly happens when outcome rates in the rate ranges at issue are generally increasing, the authors found absolute differences usually to have increased. A study by Trivedi *et al.*<sup>52</sup> relied on absolute differences between rates in examining adequacy of care (which included both treatment and control of conditions) where adequacy of care rates (especially as to treatment) were at generally high levels and increasing; and, as commonly happens in such circumstances, the authors found absolute difference between rates usually to have decreased (especially as to treatment).<sup>53</sup> A commentary<sup>54</sup> discussed the various findings and their perceived implications and stressed the need for more health disparities research and action to reduce such disparities. As was common in 2005, as it is now, neither the commentary nor any of the articles mentioned anything about the way different measures might yield different conclusions or the way any measure might be affected by general changes in the prevalence of the outcome being examined.<sup>55</sup>

---

<sup>50</sup> Vaccarino V, Rathore SS, Wenger NK, et al. Sex and racial differences in the management of acute myocardial infarction, 1994 through 2002. *N Engl J Med* 2005;353:671-682.

<sup>51</sup> Jha AK, Fisher ES, Li Z, Orav EJ, Epstein AM. Racial trends in the use of major procedures among the elderly. *N Engl J Med* 2005;353:683-691.

<sup>52</sup> Trivedi AN, Zaslavsky AM, Schneider EC, Ayanian JZ. Trends in the quality of care and racial disparities in Medicare managed care. *N Engl J Med* 2005;353:692-700.

<sup>53</sup> See [Comment on Trivedi JAMA 2006](#) (2007) regarding the authors' later effort to explain different patterns as to treatment and control, making very reasonable points in doing so, but without consideration of the generally lower rates of control compared with treatment.

<sup>54</sup> Lurie N. Health disparities – Less talk. more action. *N Engl J Med* 2005;353:727-729.

<sup>55</sup> Two studies on healthcare disparities in a recent issue of the NEJM are by groups of authors that include co-authors of the Trivedi 2005 study. See (a) Trivedi AN, Nsa W, Hausmann LRM, et al. Quality and equity of care in U.S. hospitals. *N Engl J Med* 2014;371:2298-308; and (b) Ayanian JZ, Landon BE, Newhouse JP, Zaslavsky AM. Racial and ethnic disparities among enrollees in Medicare Advantage plans. *N Engl J Med* 2014;371:2288-97.) These studies also rely on absolute differences between rates to measure healthcare disparities, again without consideration of the effects of frequency of the outcomes on the measures employed or mention that other measures might yield different conclusions.

A number of letters were published in response to the series.<sup>56</sup> Only Keppel *et al.*<sup>57</sup> directly addressed measurement issues. The letter, by three authors of the NCHS 2005 health disparities measurement monograph that recommend that all disparities be measured in terms of relative differences in adverse outcomes, presented an elaborate table to show, *inter alia*, that in four of the cases where the Trivedi study found decreasing disparities, relative differences in adverse outcomes indicated increasing disparities. The letter urged greater consistency in the reporting of disparities. Observers relying on the revised NCHS guidance, however, would instead note that relative differences in the favorable healthcare outcomes either (a) were usually consistent with the Trivedi study or (b) were usually inconsistent with the Jha study.<sup>58</sup>

The Jha study also has a role in an extreme example of the confusion among persons attempting to measure health and healthcare disparities. A 2004 *American Journal of Public Health* (AJPH) study<sup>59</sup> examined changes in racial disparities in certain fairly uncommon procedures among Medicare beneficiaries between 1986 and 1997, a period in which the procedures were generally increasing. As commonly occurs in the circumstances, the results showed that, in the main, relative difference in receipt of procedures decreased while absolute differences (and relative differences in non-receipt) increased. Relying on relative differences between rates of receiving such procedures to measure disparities, a common approach at the time, the authors found that disparities generally decreased. The Jha study mentioned above had examined changes in racial disparities in similarly uncommon procedures among Medicare beneficiaries between 1992 and 2001, also a period when rates were generally increasing. The results were much like those of the 2004 AJPH study, *i.e.*, usually decreasing relative differences in receipt of procedure and increasing absolute differences (and relative differences in non-receipt of the procedures). In this case, however, because the authors relied on absolute differences between rates to measure disparities, they found disparities usually to be increasing. A 2008 article<sup>60</sup> in *Medical Care Research and Review* then discussed the seemingly contrasting findings of directions of changes in the two studies. And without consideration of the role of choice of measure in the reportage of results or showing any awareness of the measures used in the two studies, the article opined that differing conclusions in the studies may have had to do with the absence of complete overlap of the time periods studied. In essence, the study called for more research into why studies yielded contrasting results when in fact the studies had yielded essentially the same results. See the [Spurious Contradictions](#) subpage of the [Measuring Health Disparities](#) (MHD) page of [jpscanlan.com](#).

---

<sup>56</sup> Trends in racial disparities in care. *N Engl J Med* 2005;353:2081-2084.

<sup>57</sup> Keppel KG, Percy JN, Weissman JS. Untitled. *N Engl J Med* 2005;353:2082-2083.

<sup>58</sup> I discuss the 2005 NEJM study in "Race and Mortality Revisited" (at 338), and quite a few other places. See the above-mentioned October 9, 2012 [letter](#) to Harvard University (at 34-35).

<sup>59</sup> Escarce JJ, McGuire TG. Changes in racial differences in use of medical procedures and diagnostic tests among elderly persons: 1986-1997. *Am J Public Health* 2004;94:1795-1799.

<sup>60</sup> Lê Cook B, McGuire TG, Zuvekas SH. Measuring trends in racial/ethnic health care disparities. *Med Care Res Rev.* 2009 Feb; 66(1):23-48.

An extreme example of wasted resources as result of the failure to grasp even that different measure commonly yield different conclusions about changes over time may found in AHRQ-funded review, costing between \$5 million and \$10 million, of the effects of improvement in healthcare on healthcare disparities. The study examined 4,258 studies but did not even report the measures employed in the studies. See "Race and Mortality Revisited" (at 333) and the [AHRQ's Vanderbilt Study](#) subpage of Measuring Health Disparities page.

In the prior section, I discussed that National Healthcare Disparities Reports had relied on NCHS recommendations to measure disparities in terms of relative differences in non-receipt of care. But it has not invariably implemented that approach successfully. The letter July 1, 2015 letter to the [Agency for Healthcare Research and Quality](#) discusses that confusion about measurement issues caused the 2012 National Healthcare Disparities Report to highlight as some of the largest reductions in healthcare disparities over a particular period situations where the report would also find the disparities to be substantially larger at the end of the period than at the beginning of the period. This occurred because the agency apparently measured changes in disparities according to the comparative size of the percentage point changes in each group's rate without understanding that doing so could yield a different result from the comparative size of the relative difference in the adverse outcome at the beginning of the period and the end of the period.<sup>61</sup>

As a result of contacts from AHRQ personnel, I know that AHRQ gave some attention to the issues raised in my letter. But I cannot tell whether the agency yet understands the matter. In consequence of the agency's no long publishing the tables underlying its analyses, I cannot tell precisely how it measured disparities in the 2015 report. [The Chartbook on Health Care for Blacks](#) (at 49) discusses disparities in prenatal care in terms that "Black mothers were 2.3 times as likely as White mothers to delay prenatal care to the third trimester or to not receive prenatal care at all (10.0% compared with 4.3%)." That seems to indicate that the report continues to measure healthcare disparities in terms of relative differences in adverse outcomes, notwithstanding the NCHS reversal of position. On the other hand, the [Chartbook on Healthy Living](#) presents a map ranking states according to the size of absolute differences between black and white rates of receiving early and adequate prenatal care. Given the rate ranges at issues for receipt/non-receipt of prenatal care, rankings according to relative differences in non-receipt will depart substantially from rankings according to absolute differences and general increases in appropriate care rates will tend to increase the former while reducing the latter. Thus, as in the

---

<sup>61</sup> The following may have played a part in the agency's confusion. Like NCHS, AHRQ measures changes in disparities in terms of percentage point changes in relative differences. Unlike NCHS, however, AHRQ uses the "%" sign to mean percentage points, while also using the sign to mean percent in the relative sense. See the [Percentage Points](#) subpage of the [Vignettes](#) page of [jpscanlan.com](#). In any case, the 2011 report (at 44) described changes in disparities in terms both of (a) % changes (meaning percentage point changes) in the disparities (i.e., changes in relative differences in adverse outcomes) and (b) % differences (meaning percentage point differences) between the percentage point changes in the rates of the groups being compared. In the situation where the relative difference in the adverse outcome and the absolute difference change in opposite directions, methods (a) and (b) will yield opposite conclusions as to directions of changes.

2012 report, the agency seems still to employ measures that tend to yield opposite conclusions from one another without apparent awareness of the conflict.

I have found no indication in the 2015 report that AHRQ is aware that NCHS has reversed the recommendation to rely on relative differences in non-receipt of care rather than receipt of care. One reason why AHRQ would not appreciate the significance of the revision is that AHRQ has not previously shown an understanding of the possibility that the direction of changes in disparities may turn on which relative difference is examined (though my letter should have made that clear enough). That is also a reason why AHRQ would not recognize that the NCHS reversal of position constitutes a repudiation of the first decade of National Healthcare Disparities Reports.<sup>62</sup>

In any case, confusion issues aside, the National Healthcare Disparities Reports have been universally undermined by failure to recognize the ways that the measures it employed or intended to employ tended to be affected by the prevalence of an outcome. The same holds for AHRQ funded-research.

The above discussion of the failures of understanding on the part of AHRQ should not be read to suggest that those failures are more serious than found elsewhere. The same failures exist in all parts of the federal health and healthcare disparities research establishment, including the Centers for Disease Control and Prevention (CDC), and at all institutions conducting such research.<sup>63</sup>

As indicated, the NCHS's recognition of the pattern by which the two relative differences tend to be affected by the prevalence of an outcome has done nothing to improve that agency's own research or the guidance it provides to other agencies or private researchers. Indeed, the agency's manner of dealing with such issue would tend affirmatively to lead other entities and researchers to believe that a crucial issue is a non-issue. As discussed in the 2016 JPHMP commentary, the recent guidance on which relative difference in healthcare outcomes to examine for purposes of Healthy People 2020 obscures the issue even more than prior guidance.

As discussed in the JPHMP commentary, the 2016 [article](#)<sup>64</sup> on health disparities measurement CDC and NCHS scientists/statisticians to which the commentary responded may also be deemed an affirmative obscuring of the need to address the implications of the effects of the prevalence of an outcome on measures employed in health and healthcare disparities

---

<sup>62</sup> My letter to AHRQ was written before I became aware of the NCHS reversal of position (which may not have yet then occurred).

<sup>63</sup> As indicated the NCHS's recognition of the pattern by which the two relative differences tend to be affected by the prevalence of an outcome has done nothing improve that agency's own research or the guidance it provides to other arms. Indeed, the agency's manner of dealing with such issue would tend affirmatively to lead other entities and researchers to believe that a crucial issue is a non-issue. As discussed in "The Mismeasure of Health Disparities," the recent guidance obscures the issue even more than the prior guidance.

<sup>64</sup> Penman-Aguilar A, Talih M, Huang D, et al. Measurement of health disparities, health inequities, and social determinants of health to support the advancement of health equity. J Public Health Manag Pract. 2016;22(1) (suppl):S33-S42.

research. The author's [response](#)<sup>65</sup> to the commentary, by attempting to defend the earlier work without addressing the patterns by which standard measures tend to be affected by the prevalence of an outcome or the need for researchers to attempt to sort out the effects of the prevalence of an outcome on the measure employed, constitutes a further obscuring of the issue.

\*\*\*

In the case of academic proficiency, irrespective of any change in the forces causing outcome rates of advantaged and disadvantaged groups to differ, general improvements in proficiency scores will tend to increase absolute differences between rates at which such groups reach the advanced level (where favorable outcome rates tend to be well below 50%) as reflected by movement from Row A to Row B in Table 3. On the other hand, general improvements in scores will tend to reduce such differences between rates of achieving the basic level (where favorable outcome rates are often well above 50%), as illustrated by movement from Row C to Row D. General deterioration in test performance will tend to have the opposite effect. Like patterns will tend to be observed when proficiency standards are altered or a proficiency test is replaced with one that is easier or harder.

Discussions of situations where observers commonly rely on absolute differences to measure educational disparities, but without understanding the patterns by which such measures tend to be affected by the prevalence of an outcome, may be found in the [Educational Disparities](#) page of [jpscanlan.com](#) and its [Disparities by Subject](#), [New York Proficiency Rate Disparities](#), [Education Trust High Achiever Study](#), [Education Trust Glass Ceiling Study](#), [Education Trust High Achiever Study](#), [Education Trust Glass Ceiling Study](#), [Annie E. Casey 2014 Proficiency Disparities Study](#) subpage. See also the letters to the [Annie E. Casey Foundation](#) (May 13, 2014) and [Education Trust](#) (April 30, 2014).<sup>66</sup>

But see the discussion at page 3-4 of the letter to the [Stanford Center on Poverty and Inequality](#) (Mar. 8, 2016) regarding the work of Harvard Professor Andrew Ho and Stanford Professor Sean Reardon. Professor Ho independently recognized the patterns by which absolute differences between outcome rates tend to be affected by the prevalence of an outcome and the implications of those patterns regarding the appraisal of demographic differences in proficiency rates.<sup>67</sup> While the Ho article and the illustration in its Figure 2 are focused solely on patterns by which absolute differences tend to be affected by the prevalence of an outcome, the rates that

---

<sup>65</sup> Penman-Aguilar, Talih M, Moonesinghe R, Huang M. "[Response to Scanlan Concerning: Measurement of Health Disparities, Health Inequities, and Social Determinants of Health to Support the Advancement of Health Equity](#)." *J Public Health Management Practice* 2016;22(6), 614-615 [ltr].

<sup>66</sup> In some cases, differences in proficiency outcomes (or comparable outcomes) are measured in terms of either the relative difference in the favorable outcome or the relative difference in the adverse outcome or both relative differences. Invariably, however, those employing such measures do so without apparent awareness of the ways the measures employed tend to be affected by the prevalence of the outcome. The [Harvard CRP NCLB Study](#) and [McKinsey Achievement Gap Study](#) subpages of the Discipline Disparities page and the letter to [New York City Center for Innovation through Data Intelligence](#) (June 6, 2016).

<sup>67</sup> Ho, Andrew D. 2008. "[The problem with 'proficiency': Limitations of statistics and policy under No Child Left Behind.](#)" *Educational Researcher*, 37,351–360.



underlie the figure would also form the basis for the illustrations of patterns of relative differences and odds ratios in Figures 1 and 3 in these comments. Work of Professors Ho and Reardon regarding the measurement of proficiency disparities in ways unaffected by the prevalence of an outcome,<sup>68</sup> seems to suggest approaches generally along the lines of the approach suggested in "Race and Mortality Revisited" and the October 2015 ASA letter. But, for reasons discussed in the Introduction, such matter is beyond the scope of these comments.

The following observations are in order regarding the general rigor in the study of demographic differences in terms of relative or absolute differences between outcome rates.

The attentions researchers have recent given to value judgments in the analysis of demographic differences, such as I discuss in "Race and Mortality Revisited," the 2016 JPHMP commentary, and the [Comment on Mackenbach BMJ 2016](#) (2016) as principally diverting attention from the crucial measurement issues, remains an exception to usual practice. Usual practice continues to be characterized by reliance on a preferred measure without mention of any other measure or whether such measure might, or in fact does, yield a different conclusion from the one employed in a particular study, even when the other measure may be the one most commonly employed in the circumstances. So far as I can tell, such practices is permitted in essentially all medical and other scientific journals. And it allows observers to make whatever claim they wish about how particular policies or trends disparately affect different groups. See discussions regarding the suburbanization of poverty on the [Feminization of Poverty](#) page of [jpscanlan.com](#), Table 30 (slide 118) of the University of Maryland Workshop referenced in note 18, and the letter to [Antioch Unified School District](#) (Sept. 9, 2016).

Requirements that persons analyzing demographic differences explain the reasons for their choice of measure and how other measures might or do yield contrary results would not address the key measurement issues discussed here and might even detract from them in the way recent discussions of value judgments have. But the fact that it is usually deemed permissible to use any measure one chooses to analyze demographic differences nevertheless highlights the need for fundamental reform in this area.

### **C. The Appraisal of Demographic Differences in Terms of the Proportion a Group Comprises of Persons Potentially Experiencing an Outcome and the Proportion it Comprises of Persons Actually Experiencing the Outcome**

Sometimes observers measure demographic differences on the basis of comparisons between (a) the proportion a group makes up of persons potentially experiencing a favorable or adverse outcome (the pool) and (b) the proportion it makes up of persons actually experiencing the outcome. This has long been the case with regard to selection issues in employment, where the analyses of discrimination issues commonly involve a comparison of the proportion a group makes up of the pool with the proportion is comprises of persons experiencing a favorable employment outcome. See *Hazelwood School District v United States*, 433 U.S. 299 (1977).

---

<sup>68</sup> Ho, Andrew D., and Reardon, Sean F. 2012. "[Estimating Achievement Gaps From Test Scores Reported in Ordinal 'Proficiency' Categories](#)," *Journal of Educational and Behavioral Statistics*, 37(4), 489–517.

The issues addressed here apply to those situations. See Kansas Law paper (at 23-26) and the TDHCD brief (at 23-27).

But I principally address this subject with regard to perceptions about differences between the proportion a group comprises of the pool and the proportion it comprises of persons experiencing an adverse outcome. Such comparisons are increasingly a focus of attention to disparities issues involving school discipline and criminal justice. See the [Department of Health and Human Services and Department of Education](#) (Aug. 24, 2015) and the letters to school districts and to [Texas Appleseed](#) (Apr. 7, 2015) regarding school discipline. See “[Things DoJ doesn’t know about racial disparities in Ferguson](#),” *The Hill* (Feb. 22, 2016) and [Things the President Doesn’t Know About Racial Disparities](#),” Federalist Society Blog (Aug. 5, 2016) regarding criminal justice.

Commonly such comparisons are made without indicating how one might measure the difference between (a) and (b). But a Department of Education-funded document titled “[Methods for Assessing Racial/Ethnic Disproportionality in Special Education](#)” which I discuss on the [IDEA Data Center Disproportionality Guide](#) subpage of the [Discipline Disparities](#) page, shows how to calculate both relative and absolute differences between (a) and (b) and those are the likely approaches of anyone who wants to quantify differences between (a) and (b).<sup>69</sup>

The proportion a group comprises of persons experiencing an outcome is a function of the relative difference and therefore the appraisal of disparities based on any measure of the difference between (a) and (b) is unsound in the same way the relative difference is unsound when employed without regard to the manner in which the measure tends to be affected by the prevalence of an outcome.

But there are additional reasons why one can never soundly analyze a demographic difference on the basis of a comparison of (a) and (b). The more extended treatments of measurement issues discussed in the Introduction should make clear than any sound effort to appraise a demographic difference must be based on the actual rates at which two groups experience the outcome. That also holds for observers who might maintain that the odds ratio is a measure unaffected by the prevalence of an outcome, since the actual rates are necessary in order to calculate the odds ratio. See note 49 *supra*.<sup>70</sup>

---

<sup>69</sup> In the employment setting, there occur misguided efforts to quantify the difference between (a) and (b) in terms of the statistical significance of such difference given the number of selections at issue. Problems with that approach are different from the quantification issues addressed in these comments. See the Kansas Law paper (at 26-27). But such are problems are among many issues not specifically addressed here that the Commission must address in attempting to reform analyses of demographic differences that may inform policies and programs.

<sup>70</sup> One also needs the actual rates to determine the absolute difference between rates. It is doubtful that anyone would maintain that absolute differences does not tend to be affected by the prevalence of an outcome (save in the case of uniform underlying distributions). It warrants note, however, that in cases where the only information available is (a) and (b), observers who would otherwise rely on absolute differences may well rely on some measures of the difference between (a) and (b). Comparisons of (a) and (b) will commonly yield opposite conclusions about directions of change from those based on absolute differences in matters like school discipline disparities. See the letter to [Antioch Unified School District](#) (Sept. 9, 2016) regarding a suit where the discipline disparities measure employed in the complaint tends to yield opposite conclusions from the measure commonly employed by the principal expert discussed in the complaint.

Information on (a) and (b) allow one to calculate a ratio of the rates of experiencing the outcome addressed in (b) (though not the rate ratio for the opposite outcome). See Kansas Law paper (at 24 n. 26). But such information does not enable one to divine the actual rates at which the two groups experience the outcome.

Further, the proportion a group comprises of the pool will affect both relative and absolute differences between (a) and (b) in ways that have nothing to do with the strength of the forces causing (a) and (b) to differ. That is, suppose that the rates for an adverse outcome of an advantaged group (AG) and disadvantaged group (DG) are 10% and 20% respectively in two settings (*e.g.*, different school districts). Putting aside how one might effectively quantify the difference in the circumstances of AG and DG reflected by the 10% and 20% outcome rates, there is no basis to say the difference is larger in one setting than the other. Yet, given those 10% and 20% rates, the relative and absolute differences between (a) and (b) will vary in a variety of ways depending on the proportion the disadvantaged group comprises of the pool of persons potentially experiencing the outcome.

The patterns by which the proportion a group comprises of the pool affects relative and absolute differences between (a) and (b) are discussed and illustrated in the above-mentioned my [IDEA Data Center Disproportionality Guide](#) subpage. They are also illustrated in the [slides 97 to 108](#) of the University of Maryland methods [workshop](#) mentioned in note 9.

It is unnecessary to illustrate those patterns here, however. For the previously mentioned issues provide sufficient reason why demographic differences should never be analyzed on the basis of comparisons between the proportion a group comprises of the pool and the proportion it comprises of persons experiencing some favorable or adverse outcome.

The failure of the government or research communities to recognize these issues does, however, further illustrate the need the Commission to take actions of the kind outlined in Part III.

#### **D. Subgroup Effects Issues**

Section A discusses the failure of persons examining subgroup effects or drawing inferences about processes on the basis of the comparison of relative effects of a factor on different subgroups (or in different settings) to understand the reasons to expect a factor affecting an outcome rate to cause a larger proportionate change in the rate for the group with the lower baseline rate for the outcome while causing a larger proportionate change in the opposite outcome rate for the other group.<sup>71</sup> That section and the materials it references also explain that, so far as the published reveals, never has anyone identifying a subgroup effect or drawing an inference based on the comparative sizes of relative effects recognized that examination of the opposite outcome could (or commonly would or in fact would in the particular situation

---

<sup>71</sup> See also the October 9, 2012 letter to [Harvard University](#) at 39-40.

examined) form the basis for an opposite subgroup effect or opposite inference. See especially "Race and Mortality Revisited" at 339-341.

Irrespective of such pattern, however, it should be recognized that it is illogical to believe that, absent a subgroup effect as such effect is typically analyzed, one would expect a factor that affects on outcome rate to have the same relative effect on different baseline rates for the outcome. For if a factor causes two groups to experience equal proportionate changes in different baseline rates for one outcome, it necessarily causes the groups to experience unequal proportionate changes in their rates for the opposite outcome. Since there is no more reason to expect equal proportionate changes in one outcome than there is to expect equal proportionate changes in the opposite outcome, it is illogical to expect equal proportionate changes in either outcome.

Aspects of this issue are discussed on following subpages to the [Scanlan's Rule](#) page of [jpscanlan.com](#). The pages are in many respect duplicative of one other. But in my view (at least of the time of creating it) each added an additional element of sufficient importance, or caused a particular aspect of the matter to be sufficiently better focused, to warrant the creation of the page.

The [Subgroups Effects](#) subpage discusses this issue with regard to the clinical setting, while explaining the implications of the above-described pattern of relative effects with regard to employing an observed risk reduction in a clinical trial to estimate the absolute risk reduction (and corresponding number-needed-to-treat) in circumstances involving baseline rates different from that in the trial. Illustrations pertinent to this subpage may be found in my presentation at the 2009 Joint Statistical Meetings "[Interpreting Differential Effects in Light of Fundamental Statistical Tendencies](#)" ([abstract](#))

The [Subgroup Effects – Nonclinical](#) subpage discusses the subject with regard to perceptions about subgroup effects in non-clinical settings.

The [Illogical Premises](#) subpage discusses the illogic of the belief that a factor will tend to cause the same proportionate effect on different baseline rates while addressing whether similar expectations for the absolute difference and odds ratio would be illogical as well.

The [Illogical Premises II](#) discusses that, for the same reason that it is illogical to expect a factor typically to show the same relative effect on different baseline rates of experiencing the outcome, it is illogical to regard a particular rate ratio as reflecting the same strength of an association for different baseline rates for an outcome.

The [Inevitability of Interaction](#) subpage discusses that any time a factor in fact shows equal proportionate changes for rates of experiencing an outcome (thus, the absence of an interactive effect) it will necessarily show a different proportionate effect on the opposite outcome.

The [Interactions by Age](#) subpage, which is closely related to the [Life Tables Illustrations](#) subpage, discusses the fact that almost invariably in comparisons of age groups with

substantially different mortality rates one will find opposite patterns of interaction depending on whether one examines a factor's effect on mortality or on survival (that is, that the age group with the smaller proportionate effect on its mortality rate will show the larger proportionate effect on its survival rate).

Other treatments of these issues may be found in my [Comment on FDA Proposed Subgroup Regulations](#) (May 16, 2014) [Comment on European Medicines Agency Subgroup Guidelines](#) (July 31, 2014).

Relatively succinct treatments of the issues with reference to journal articles regarding subgroup effects, reporting heterogeneity, perceptions about the constancy of the rate ratio across different baseline rates, calculation of number-needed-to-treat, or stratified medicine may be found in [Comment on Hingorani BMJ 2013](#), [Comment on Barrett CMAJ 2004](#) (2012), [Comment on Delpierre BMC Public Health 2012](#) (2012), [Comment on Wang Emerging Themes in Epidemiology 2009](#) (2012), [Comment on Altman BMJ 2003](#) (2011), [Comment on Kent Trials 2010](#) (2011), [Comment on White BMC Med Res Meth 2005](#) (2011), , [Comment on Schwartz BMJ 2007](#) (2011), [Comment on Cook BMJ 1995](#) (2011), [Comment on Huisman Int J Epidemiol 2007](#) (2010), [Comment on Chatellier BMJ 1996](#) (2011), [Comment on Mullins Health Affairs 2010](#) (2010), [Comment on Berrington de Gonzalez MEJM 2010](#) (2010), [Comment on Sun BMJ 2010](#) (2010), [Comment on Gabler Trials 2009](#) (2009), [Comment on Kawachi Health Affairs 2005](#) (2008), [Comment on Gan NEJM 2000](#) (2007), [Comment on Mustard JECH 2003](#) (2007), [Comment on Thurston AJE 2005](#) (2007), [Comment on Martikainen JECH 2007](#) (2007), [Comment on Kaplan JECH 2006](#) (2006).

As discussed in the introduction to Part I, the patterns discussed in this section are merely manifestation of other patterns described in the introduction and Section A. But given the resources devoted to studies of subgroups effects in clinical and other contexts, and the varied flawed inferences based on perceptions about such effects, perceptions about subgroups effects warrant special attention from the Commission.

## **II. The Fundamental Unsoundness of Analyses of Discrimination Issues That Examine Data Solely on Persons Who Accepted Some Outcome or Situation.**

Most discrimination cases that have yielded recoveries approaching or exceeding \$100 million have been based on analyses that examined data solely on persons who accepted some outcome or situation. In the 1990s such cases principally involved claims that persons who were hired into grocery stores or other retailers were disproportionately assigned to jobs that were believed to be less desirable than other jobs. See, e.g. "[Multimillion-Dollar Settlements May Cause Employers to Avoid Hiring Women and Minorities for Less Desirable Jobs to Improve the Statistical Picture](#)," *National Law Journal* (Mar. 27, 1995) and "[Unlucky Stores: Are They All Guilty of Discrimination?](#)" *San Francisco Daily Journal* (Jan. 29, 1993). More recently, such cases have involve claims that lenders disproportionately assigned loans to minorities to subprime status, as in the cases brought by the Department of Justice against Countrywide Financial Corp. and Wells Fargo Bank that were settled for \$335 and \$175 million. See "[The Perverse Enforcement of Fair Lending Laws](#)," *Mortgage Banking* (May 2014), and "[Fair Lending Studies Paint Incomplete Picture](#)," *American Banker* (April 24, 2013).

I first discussed this issue with regard so-called job segregation or assignment discrimination claims in "[Illusions of Job Segregation](#)," *Public Interest* (Fall 1988) and "[Are Bias Statistics Nonsense?](#)" *Legal Times* (Apr. 17, 1989). The former item presents a number of tabular illustrations of why one cannot draw inferences about discriminatory exclusion from putatively better jobs without consideration of the part of the applicant pool that was not interested in the putatively poorer job.

Table IV of that article explains why one also cannot draw inference about the absence of discrimination of the basis of weaker qualifications of persons from a particular group among persons who are hired, which has been asserted as showing that the group is being favored. The reason one cannot draw such an inference is that the group with weaker qualifications among applicants will commonly also have weaker qualifications among persons who are hired even when the group has been subject to discrimination. I have explained this point with regard to arguments that weaker qualifications among athletes from a particular group or college-admissions from a particular group suggests that such groups are being favored "[The Perils of Provocative Statistics](#)," *Public Interest* (Winter, 1991). I have also explained it with regard to the claim that higher default rates among minority borrowers suggests that they are being favored in "[Both Sides Misuse Data in the Credit Discrimination Debate](#)," *American Banker* (July 22, 1998).

The job segregation/assignment discrimination claims discussed in "Illusions of Job Segregation" or other articles of the 1980s or 1990s involved binary outcomes (at least in the sense that the putative victims of discrimination were being disproportionately assigned to a less desirable rather than a more desirable situation). The same holds for claims of discriminatory assignment of loans to subprime status in recent fair lending cases. But as discussed in the Kansas Law paper (Section F, at 32-35) and the TDHCD brief (Section I.C, at 27-30), the fundamental failing of analyses that fail to examine data on the entire universe at issue also applies to situations of continuous variables like salaries and loan terms.

That is, the validity of allegations that persons from a particular group received lower starting salaries or higher interest rates than persons another group cannot be appraised without consideration of persons who refused to accept those terms (and persons offered no terms at all). Further, in the employment context, whether one group's salary progression was impeded by discrimination cannot be analyzed without consideration of the treatment of the persons who left the employer because of what such persons regarded as inadequate salary progression or for any other reasons.

But an understanding of the implications of the failure to examine data on the entire universe at issue is as universally absent from analyses of discrimination issues as the understanding of patterns by which measures tend to be affected by the prevalence of an outcome is from analyses of demographic differences in outcome rates.<sup>72</sup>

---

<sup>72</sup> The aspect of risk distributions underlying the patterns discussed in part are implicated in all efforts to analyze discrimination issues in the same way they are implicated in all efforts to adjust for pertinent factors in any analysis. See "[The Perils of Provocative Statistics](#)," *Public Interest* (Winter, 1991) and the [Sears Case Illustration](#) subpage. In my view rarely if ever are efforts to adjust for differences in characteristics fully adjust for such differences.

### III. Recommendations

Set out below are a number of recommendations for Commission action to address the issues discussed in Parts I and II and related matters.

First, the Commission should create a committee to explore problems in analyses involving demographic and other differences in outcome rates (including with regard to interpretations of subgroup effects in clinical and non-clinical settings) and to recommend actions to improve such analyses. In accomplishing its tasks, the committee should become familiar with all the materials referenced above and seek informed input from the scientific community.

Second, the Commission should recommend that Congress establish a permanent body charged with appraising the soundness of statistical analyses in matters affecting public policy, including law enforcement. It might be suggested that there already exist such bodies, including the National Academies and the Government Accountability Office. But these and like entities have failed even to recognize that reducing adverse outcomes, whether involving health and healthcare, lending, school discipline, school performance, criminal justice, or any other matter, tends to increase rather than reduce relative differences in rates of experiencing those outcomes. Whether or not a creation of a new entity is necessarily a solution, some action is necessary to ensure that statistical analyses involving important policy issues receive much greater scrutiny than they have in the past.

Third, the Commission should recommend to Congress that it require reports of the agencies involved with funding or conducting statistical analyses involving outcome rates, or involved with law enforcements based on such analyses, describing actions they will take to address the issues discussed in these comments and the materials they reference.

Fourth, the Commission should recommend that Congress require that requests for federal funding of health and healthcare disparities and other research involving differences in outcome rates include statements like the following:

1. We are aware that there exist patterns by which measures commonly employed in this type of research tend to be affected by changes in the prevalence of the outcome examined irrespective of (a) actual changes in differences in the circumstances of advantaged and disadvantaged groups or (b) effects of policies aimed at mitigating those differences.

---

Moreover, observers seem rarely to acknowledge even the obvious shortcoming of such efforts. But, while the failure to adjust for characteristics is commonly a serious problem in analyses of discrimination issues and many other issues, such failure does raise the same issues as to the fundamental unsoundness of analyses of demographic differences as the failure to recognize the ways measures tend to be affected by the prevalence discussed in Part I and the failure to examine the entire universe at issue discussed in Part II.

2. We intend to attempt to distinguish between the effects of the patterns by which measures tend to be affected by the prevalence of an outcome and (a) and (b) in the following manner:

Studies themselves should include at their beginning the following statements:

1. This study has (or has not) attempted to distinguish between the effects of changes in the prevalence of an outcome on the measures employed and (a) actual changes in differences in the circumstances of advantaged and disadvantaged groups (b) effects of policies aimed at mitigating those differences.

2. Because of 1, this study may be (should not be) used to inform policy.

The language can be adjusted to address the situation of research aimed at appraising the effects of factors other than policies on demographic differences. It can also be adjusted to address situations where, rather than examining changes in differences over time, researchers compare the size of some difference within settings differentiated other than temporally (and where the settings differ in the overall prevalence of an outcome).

Similar statements should be required for the continuation of funding already authorized.

Fifth, the Commission should recommend that Congress take all steps necessary to ensure that no federal law enforcement actions are based the belief that reducing the frequency of an adverse outcome tends to increase relative demographic differences in rates of experiencing the outcome or the proportion disadvantaged groups make up of persons experiencing those outcomes.

Sixth, the Commission should recommend that Congress identify all existing legislation (a) that reflects the belief that reducing the frequency of an adverse outcome will tend to reduce relative demographic differences in rates of experiencing the outcome or the proportion disadvantaged groups make up of persons experiencing the outcome; (b) that require the monitoring of demographic differences with regard to some outcome; (c) that impose liability for a practice that has a disparate impact; (d) that require implementation of a less discriminatory alternative to practices having a disparate impact. Congress should then consider options for eliminating any false beliefs reflected in such legislation and for either clarifying how differences and disparate impacts are to be measured or eliminating the requirements.

Seventh, the Commission should recommend that Congress require that federal agencies take the same actions regarding regulations that the prior paragraph suggests Congress take regarding legislation.



## **Corrections to Version of Comments Submitted November 14, 2016**

### **Corrections made on November 17, 2016:**

Page 2, last full paragraph last line: “increased racial/ethnic” changed to “increased relative racial/ethnic”

Page 6, note 8, line 2: “individual” changed to “individuals”

Page 9, second paragraph, lines 2-3: “between circumstances” changed to “between the circumstances”

Page 9, third paragraph, line 3: “by the” changed to “by their”

Page 10, note 17, line 5: “quantifications” changed to “quantification”

Page 11, third full paragraph, line 1: “subgroup analyses” changed to “subgroup effects”

Page 15, first full paragraph, line 4: “pass rates” changed to “failure rates”

Page 17, last paragraph, line 1: “large part of those communities expect” changed to “large proportion of persons analyzing demographic differences expects”

Page 17, last paragraph, line 2: “large part of the community” changed to “large proportion of such persons”

Page 18, second full paragraph, line 6: “cause more poorer” changed to “cause poorer”

### **Corrections made on November 18, 2016:**

Page 23, second paragraph, line 3: “to test score” changed to “to the test score”

Page 24, second paragraph, last line: “might be actually” changed to “might actually”

Page 24, third paragraph, line 6: “to the NCHS” changed to “to NCHS”

Page 24, fourth paragraph, first line: “led their” changed to “led to their”

Page 25, third full paragraph, first line: “has now reversed” to “has recently reversed”

Page 27, second line: “reducing the reducing public” changed to “reducing public”

Page 28, second full paragraph, line 8: “Meanwhile it” changed to “Meanwhile, it”

Page 28, second full paragraph, fourth last line: “Disparate Impact.” Changed to “Disparate Impact page.”

Page 28, second full paragraph, third last line: “Impact and each of its subpage” changed to “Impact page and each of its subpages”

Page 28, last paragraph, first line: “legislations impose” changed to “laws impose”

Page 29, paragraph from prior page, third last line: “government” changed to “governments”

Page 36, first paragraph, second line: “measure commonly” changed to “measures commonly”

Page 36, first paragraph, line 4: “improvement” changed to “improvements”

Page 36, third paragraph, “Healthcare” changed to “Health Care” (and link added)

Page 36, third paragraph, line 9: “Chartbook on Health Living” changed to “Chartbook on Healthy Living”

Page 37, note 63, line 4: “researchers that” changed to “researchers to believe that”

Page 38, third full paragraph, second line: “the work of work of” changed to “the work of”

Page 38, third paragraph beginning on the page, line 6: “is focused” changed to “are focused”

Page 39, paragraph from prior page, third last line” “outcome, seems to be suggest” changed to “outcome seems to suggest”

Page 39, second full paragraph, last line: comma added after “18”

Page 39, third full paragraph, second last line: “nevertheless highlight” changed to “nevertheless highlights”

**Corrections made on November 23, 2016:**

Page 39, Section C, first paragraph, first line: “observer measure” changed to “observers measure”

Page 39, Section C, first paragraph, line 2: “the (a)” changed to “(a) the”

Page 39, Section C, first paragraph, line 3: “outcomes outcome” changed to “outcome”

Page 42, second full paragraph, last line: “matter sufficiently be” changed to “matter to be sufficiently”

Page 42, paragraph beginning “The Inevitably..”, first line: ‘subpage why’ changed to “subpage discusses that”

Page 43, paragraph from prior page beginning “The Interactions...”, fourth last line: “rates, one” changed to “rates one”

Page 44, first full paragraph, line 3: “persons are hired” changed to “persons who are hired”

Page 44, last paragraph, line 2: “at issues” changed to “at issue”