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

November 12, 2015

ELECTRONICALLY TRANSMITTED

Inez H. Friedman-Boyce, Esq., Co-Chair
Lisa Pirozzolo, Esq., Co-Chair
Members of the Board of Directors
The Lawyers' Committee for Civil Rights
  and Economic Justice
294 Washington Street, Suite 443
Boston, MA 02108

> Re: Problems in Lawyers' Committee Analyses of Demographic
> Differences in Outcome Rates

 Dear Co-Chairs Friedman-Boyce and Pirozzolo and Board Members:

	The purpose of this letter is to bring to the attention of the Lawyers' Committee for Civil Rights and Economic Justice problems in its analyses of demographic differences arising from the failure to understand certain things about the ways measures of differences between outcome rates tend to be affected by the frequency of an outcome, including that reducing the frequency of an outcome like suspension from school tends to increase, not decrease, (a) relative demographic differences in rates of experiencing the outcome and (b) the proportions groups most susceptible to the outcome make up of persons experiencing the outcome. While issues of the type addressed in this letter pertain to many Lawyers' Committee activities, the particular focus of this letter is a November 2014 Lawyers' Committee report titled "Not Measuring Up: The State of School Discipline in Massachusetts."[1] I will be using data from the report to illustrate certain things about the misunderstanding of data on demographic differences, both generally and with respect to matters of special relevance to places like Massachusetts, in a seminar titled "The Mismeasure of Health Disparities in Massachusetts and Less Affluent Places" at the Quantitative Health Sciences Department of the University of Massachusetts Medical School on Wednesday, November 18, 2015. A flyer for the seminar may be found here and an abstract may be found here.

---

[1] To facilitate consideration of issues raised in letters such as this I include links to referenced materials in electronic copies of the letters. Electronic copies of the letter may be found on the Institutional Correspondence subpage of the Measuring Health Disparities page of jpscanlan.com. If the online copy of the letter is corrected subsequent to its initial transmission, such fact will be noted on the online copy. The online copy of this letter may also be accessed by this link.

On occasion I write to institutions or organizations whose activities involve the interpretation of data on demographic differences in the law or the social or medical sciences alerting them to ways in which their interpretations are undermined by the failure to understand patterns by which standard measures of differences between favorable or adverse outcome rates of advantaged and disadvantaged groups – or differences between the proportion a group makes up of persons potentially experiencing an outcome and the proportion it makes up of persons actually experiencing the outcome – tend to be systematically affected by the overall frequency of an outcome. Recipients of other letters involving interpretive issues of the type discussed in this letter include Robert Wood Johnson Foundation (Apr. 8, 2009), National Quality Forum (Oct. 22, 2009), Institute of Medicine (June 1, 2010), The Commonwealth Fund (June 1, 2010), United States Department of Education (Apr. 18, 2012), United States Department of Justice (Apr. 23, 2012), Board of Governors of the Federal Reserve System (Mar. 4, 2013), Harvard University (Oct. 9, 2012), Harvard Medical School, Massachusetts General Hospital, et al. (Oct. 26, 2012), Senate Committee on Health, Education, Labor and Pensions (Apr. 1, 2013), Mailman School of Public Health of Columbia University (May 24, 2013), Investigations and Oversight Subcommittee of House Finance Committee (Dec. 4, 2013), Education Trust (April 30, 2014), Annie E. Casey Foundation (May 13, 2014), Institute of Medicine II (May 28, 2014), IDEA Data Center (Aug. 11, 2014), Education Law Center (Aug. 14, 2014), Financial Markets and Community Investment Program, Government Accountability Office (Sept. 9, 2014), Wisconsin Council on Families and Children's Race to Equity Project (Dec. 23, 2014), Portland, Oregon Board of Education (Feb. 25, 2015), Vermont Senate Committee on Education (Feb. 26, 2015), United States Department of Justice and City of Ferguson, Missouri (Mar. 9, 2015), Senate Committee on Health, Education, Labor and Pensions II (Mar. 20, 2015), Texas Appleseed (Apr. 7, 2015), City of Minneapolis, Minnesota (June 8, 2015), Agency for Healthcare Research and Quality (July 1, 2015), Department of Health and Human Services and Department of Education (Aug. 24, 2015), McKinney, Texas Independent School District (Aug. 31, 2015), Chief Data Scientist of the Office of Science and Technology Policy (Sept. 8, 2015), American Statistical Association (Oct. 8, 2015), and House Judiciary Committee (Oct. 19, 2015). An *amicus curiae* brief I filed on November 17, 2014, in *Texas Department of Housing and Community Development, et al. v. The Inclusive Communities Project, Inc*., Sup. Ct. No. 13-1371 (TDHCD brief), might be deemed a similar communication to the United States Supreme Court.

All of these letters address, usually among other things, the failure of the recipient entities to understand the pattern whereby the rarer an outcome the greater tends to be the relative difference in experiencing it and the smaller tends to be the relative difference in avoiding the outcome. In the instant context, such pattern means that relaxing any standard and thereby reducing the frequency of an adverse outcome, while tending to reducing relative differences between rates at which advantaged and disadvantaged groups meet the standard (and avoid the adverse outcome), tend to increase relative differences in failing to meet the standard (and experience the adverse outcome). Fairly succinct, recent explanations of why this occurs may be found in my "Things government doesn't know about racial disparities," T*he Hill* (Jan. 28, 2014);"The Paradox of Lowering Standards," *Baltimore Sun* (Aug. 5, 2013); "Misunderstanding of Statistics Leads to Misguided Law Enforcement Policies," *Amstat News* (Dec. 2012); "'Disparate Impact': Regulators Need a Lesson in Statistics, *American Banker* (June 5, 2012), "The Lending Industry's Conundrum," *National Law Journal* (Apr. 2, 2012).

These items also explain the anomalies in federal civil rights enforcement arising from the government's failure to understand these patterns. In particular, for some time the government has encouraged lenders to relax lending standards and public schools to relax discipline standards in order to reduce relative racial differences in rates of experiencing adverse borrower outcomes and adverse discipline outcomes. Unaware that relaxing standards tends to increase, not reduce, relative differences in failure to meet the standards, the government continues to monitor the fairness of lending and discipline practices on the basis of relative differences in adverse borrower and discipline outcomes. Thus, by acceding to government encouragements to relax standards lenders and public schools increase the chances that the government will accuse them of discrimination.

Recent, more extensive treatments of those anomalies and related matters may be found in my "Race and Mortality Revisited," *Society* (July/Aug. 2014); "The Perverse Enforcement of Fair Lending Laws," *Mortgage Banking* (May 2014); "Measuring Health and Healthcare Disparities," 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 the above-referenced *amicus curiae* brief and letter to the American Statistical Association. [2]

The extended treatments contain a number of graphical and tabular illustrations of the referenced pattern and other patterns by which measures tend to be affected by the frequency of an outcome. Many more graphical and tabular illustrations may be found in methods workshops given in recent years at American universities.[3] A simple tabular illustration is set out in Table 1 below.

The table, which reflects the same hypothetical used in the recent shorter articles as well as several of the extended treatments, 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

---

[2] Older articles with a focus on implications of the failure to the above-described pattern of relative differences in legal settings include: "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); An Issue of Numbers," *National Law Journal* (Mar. 5, 1990).

[3] 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).

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 first pair of shaded columns, with tan shading in the electronic copy of the letter).[4]   Based on a situation where AG and DG each make up half of the test takers, the final columns (shaded red in the electronic copy of the letter) show the proportion DG makes up of those who pass and those 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 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 was reduced in frequency (test failure) increases, while the relative difference in the opposite outcome (test passage, which increased in frequency) declines.

It warrants note at this point that the fact that lowering a test cutoff tends to reduce relative differences in pass rates is well known, especially among lawyers who deal with employment discrimination issues. Such fact underlies the universal belief that lowering test cutoffs tends to reduce the disparate impact of employment and other tests where some groups outperform others as well as the requirement that employers justify how high they have set a test

---

[4] While I commonly refer to patterns of relative differences in this letter, 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. 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 relative differences tend to be affected by the frequency 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 some earlier works, 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) yielding a rate ratio for the favorable outcome that was below 1. 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 frequency of the outcome.

cutoff.[5] But even though the fact that lowering a cutoff will tend to increase relative differences in failure rates is implied in the widely known fact that lowering the cutoff tends to reduce relative differences in pass rates (see note 14 of the American Statistical Association letter) at least so far as the published record reveals, the fact that lowering a cutoff tends to increase relative differences in failure rates is virtually unknown. That seems to be the case even among agencies like the Departments of Justice and Education that have been dealing with issues concerning demographic differences in testing outcomes for decades.

As shown in the final two columns of Table 1, 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 unaffected by the cutoff, lowering the cutoff would increase both the relative difference and the absolute difference between the proportion DG makes up of test takers and the proportion it makes up of persons who fail (while reducing both the relative difference and the absolute difference between the proportion DG makes up of test takers and the proportion it makes up of persons who pass).

As with the fact that reducing the frequency of an adverse outcome tends to increase relative differences in rates of experiencing the outcome, the fact that reducing the frequency of an adverse outcome tends to increase the proportion disadvantaged groups make up of persons experiencing an outcome is also almost universally unknown. See the March 9, 2015 letter to the Department of Justice (DOJ) and the City of Ferguson, Missouri explaining that findings in the DOJ's March 4, 2015 report titled "Investigation of the Ferguson Police Department" that police and court procedures of Ferguson, Missouri had a disparate impact on the city's African American residents is based on the mistaken premise that reducing the frequency of adverse interactions between the police/courts and the city's residents would tend to reduce the proportion African Americans make up of persons subject to those interactions. See also the August 24, 2015 letter to the Department of Health and Human Services and Department of Education explaining to those agencies that a policy statement the agencies jointly issued on preschool discipline practices is based on the mistaken premise that generally reducing preschool suspensions and expulsions will tend to reduce the proportions disadvantaged groups make up of persons experiencing those outcomes. In both cases the premises are the opposite of reality.

The above-described pattern of relative differences can also be illustrated with data from the recent Lawyers' Committee report. Table 2 below, which also appears as Table 1 of the Massachusetts Disparities subpage of the Discipline Disparities page of jpscanlan.com, is based on data underlying the report's observation that racial disparities in out-of-school suspensions were slightly larger in Massachusetts than nationally and that disparities between out-of-school suspension rates of students with disabilities and students without disabilities were much larger in Massachusetts than nationally. As explained above, because suspension rates are lower in Massachusetts than nationally, relative differences in suspension rates will tend to be larger in

---

[5] Whether relaxing a standard in fact reduces the standard's impact, properly measured, is a complex subject. See Section E (at 27-32) of the Kansas Law paper mentioned above.
.

Massachusetts than nationally, while relative differences in rates of avoiding suspension will tend to be smaller in Massachusetts than nationally.

In Table 2, the "Type" column shows whether the comparison is between whites and blacks or between students without disabilities and students with disabilities. For the white-black comparison the advantaged groups (AG) and disadvantaged groups (DG) are whites and blacks; for the Gen-Ed – Special Ed comparisons AG and DG are students without disabilities and students with disabilities. The column "DG/AG Ratio Susp" shows the ratios the DG suspension rates to the AG suspension rates. These are the figure on which the report's statements about the comparative size of disparities in Massachusetts and nationally are based.

The column "AG/DG Ratio No Susp" shows ratios of AG's rate of avoiding suspension to DG's rate of avoiding suspension." Together the two ratio columns show the common pattern whereby the setting where the adverse outcome is less common shows the larger relative difference in the adverse outcome, but the smaller relative difference in the corresponding favorable outcome, than the setting where the adverse outcome is more common.

The "EES" (for "estimated effect size") column contains a measure that, as discussed in "Race and Mortality Revisited," is the most plausible indicator of the strength of the forces causing outcome rates of AG and DG to differ. The measure indicates that the strength of those forces is less in Massachusetts than nationally. The table can be compared with Table 8 or "Race and Mortality Revisited" (at 342), which presents data on racial differences in suspension rates in preschool and K-12, though in the case of the preschool/K-12 comparison, the EES figures were approximately equal.

Table 2: Out-of-school suspension rates for AG and DG (as identified in the text), in Massachusetts and nationally, with measures of difference

Type	Area	AG	DG	DG/AG Ratio-Susp	AG/DG Ratio - No Susp	EES
White-Black	Massachusetts	2.7%	10.0%	3.70	1.08	0.65
White-Black	National	4.6%	16.4%	3.57	1.14	0.71
Gen Ed - Spec Ed	Massachusetts	2.8%	8.5%	3.04	1.06	0.43
Gen Ed - Spec Ed	National	6.0%	13.0%	2.17	1.08	0.58

The Lawyers' Committee report makes several statements about the fact that the proportion certain groups make up of persons suspended is so much larger than the proportion such groups make up of enrolled students. Analyses of disparities issues based on the proportion a group makes up of persons potentially experiencing an outcome and the proportion it makes up of persons actually experiencing an outcome are problematic for reasons beyond the central point of this letter. See Section I.B of the amicus curiae brief, the IDEA Data Center Disproportionality Guide and Disabilities – Public Law 104-446 subpages of the Discipline Disparities page of jpscanlan.com, and slide 97 to 108 of the Maryland methods workshop identified in note 3 *supra*. Further, in most cases, the report does not present information on the proportion each group makes up of enrollment nationally.

In discussing (at 11) that in Massachusetts male students, while "just under 52% of enrollment, were involved in 71% of all disciplinary removals, slightly higher than the national average," the report provides information in a footnote on national figures. The national figures in the footnote, however, show separately the proportion male students made up of persons suspended once and the proportion they made up of persons suspended more than one once (and not the proportion they made up of persons suspended one or more times). But it reasonable to expect that the difference between the proportion male students make up of students (which will differ little between Massachusetts and the nation) and the proportion they make up of persons suspended one or more times will be larger, both in relative and absolute terms, in Massachusetts than nationally, simply because suspension rates are so much lower in Massachusetts than nationally.

Other materials pertinent to the suspension issue include the subpages to the Discipline Disparities page discussing data from across the country showing that recent general reductions in discipline rates have been accompanied by increased relative racial/ethnic differences in discipline rates. See the following subpages pertaining to patterns in the jurisdictions identified in the names of the subpages: Los Angeles SWPBS, Denver Disparities, Florida Disparities, Maryland Disparities, California Disparities, Connecticut Disparities, Maryland Disparities, Minnesota Disparities, Rhode Island Disparities, St. Paul Disparities, Minneapolis Disparities, Beaverton (OR) Disparities, Portland (OR) Disparities, Montgomery County (MD) Disparities, and Henrico County (VA) Disparities.

See also the DOE Equity Report subpage of the Discipline Disparities page. [6] That subpage discusses data in the DOE's November 2012 report titled "Helping to Ensure Equal Access to Education: Report to the President and Secretary" showing that, notwithstanding claims of the Department of Education and Justice that zero tolerance policies are responsible for large relative racial differences in adverse discipline outcomes, relative racial differences in expulsion rates are smaller in districts with zero tolerance policies than in districts without such policies.

The report (at A2) commends the Massachusetts Department of Elementary and Secondary Education for collecting certain types of data prior to implementation of a new law requiring collection of data because of the value of the earlier data in serving as a baseline against which to appraise the impact of a new law. For reasons discussed above, however, it is essential to recognize that one cannot soundly appraise the new law's impact on demographic differences in disciplinary actions without understanding that general reductions in discipline actions will tend to increase relative differences in discipline rates, as well as increase the proportion disadvantaged groups make up of persons disciplined, irrespective of any change in the strength of the forces causing discipline rates of advantaged and disadvantaged groups to differ.

---

[6] 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, and Vignettes. The pages have close to a hundred subpages.

More generally, as discussed in the extended treatments of these issues on page 3, it is not possible to appraise the strength of the forces causing outcome rates of advantaged and disadvantaged groups to differ – or to draw inferences about the nature of such forces on the basis of the comparative size of relative differences (either as to the favorable or adverse outcome) as to differing types of outcomes or in different settings – without fully understanding the patterns described in those treatments.

Further, other standard measures of differences between outcome rates tend also to be systematically affected by the frequency of an outcome. Thus, the issues addressed in those treatments are likely to pertain to wide range of activities of the Lawyers' Committee involving the analyses of demographic differences, regardless of the particular measure employed.

Sincerely,

/s/ James P. Scanlan

James P. Scanlan