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

April 23, 2012

The Honorable Eric H. Holder, Jr.
Attorney General
The Honorable Thomas E. Perez
Assistant Attorney General for the Civil Rights Division
United States Department of Justice
950 Pennsylvania Avenue, NW
Washington, DC 20530-0001

>   Re: The Department of Justice's Misunderstanding of Statistical Concepts Relating to
>   the Enforcement of Civil Rights Law

Dear Attorney General Holder and Assistant Attorney General Perez:

On occasion I write to institutions whose missions involve the interpretation of data on
demographic differences in the law and the social and medical sciences, alerting those
institutions to problems in their interpretations arising from the failure to recognize the ways that
standard measures of differences between rates of experiencing favorable or adverse outcomes
tend to be affected by the overall prevalence of an outcome. This letter is prompted by media
attention to matters involving racial differences in public school discipline and mortgage lending
where the Department of Justice's interpretation of the relationships between the policies at issue
and the observed racial differences is the opposite of reality and where policies encouraged by
the Department will tend to increase the differences prompting the Department's concern in the
matters. The Department's misunderstanding of these relationships is addressed Section A
below. Certain related matters are addressed in Section B. Certain technical issues are
addressed in Section C.

**A. The Relationship Between the Frequency of Adverse Outcomes and the Size of Racial
Differences in Experiencing those Outcomes**

In March 2012 the Department of Education's Office of Civil Rights released data showing
large racial differences in public school discipline rates. Substantial media coverage of the
release of the data attributed the size of the disparities to stringent discipline policies in effect in
recent decades. A similar perception apparently caused the Department of Justice in July 2011 to
join with the Department of Education in creating the Supportive School Discipline Initiative to

promote the exploration of more lenient alternatives to existing discipline policies. Large racial differences in discipline rates, however, are not the consequence of stringent discipline standards. As explained below, the more stringent the standard, the smaller will tend to be racial differences in discipline rates.

In December 2011, the Department announced the $335 million dollar settlement of claims of mortgage lending discrimination in *United States v. Countrywide Financial Corporation* et al., No. CV11-10540 (C.D. Cal). The Department's complaint in the case suggests both (a) that lenders identifying substantial racial differences in adverse lending outcomes will be expected to implement less discriminatory alternatives and (b) that reducing the frequency of adverse lending outcomes will result in smaller racial differences in rates of experiencing those outcomes. As with the Department's perception about the association between large relative differences in public school discipline rates and stringent discipline standards, the Department's perception that reducing the frequency of adverse lending outcomes will tend to reduce racial differences in experiencing those outcomes is mistaken. The opposite is the case.

Inherent in the shapes of distributions of factors associated with experiencing an outcome is a 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 it. Links to over 160 references explaining this pattern and related patterns of relationships between the prevalence of an outcome and measures of differences between outcome rates in particular settings may be found on the Measuring Health Disparities[1] page (MHD) of jpscanlan.com. Nuances of the patterns are described on the Scanlan's Rule page (SR) of the same site. A number of published articles explaining various implications of these patterns and the failure to understand them, the most recent of which discusses the *Countrywide* case, may be found in the margin.[2]

The patterns whereby relative differences in experiencing an outcome and relative differences in avoiding the outcome are related to the prevalence of an outcome, including the pattern whereby

---

[1] To facilitate consideration of the issues raised in letters such as this I make available electronic copies of the letters on the Institutional Correspondence sub-page of the Measuring Health Disparities page of jpscanlan.com. Underlinings in this letter reflect links to the underlined material in an electronic copy of the letter maintained on the referenced Institutional Correspondence sub-page.

[2] "The Lending Industry's Conundrum," *National Law Journal,* Apr. 2, 2012; "Can We Actually Measure Health Disparities?," *Chance* 2006;19(2) :47-51; "Race and Mortality," *Society* 2000;37(2):19-35 (reprinted in *Current* 2000 (Feb)); "Both Sides Misuse Data in the Credit Discrimination Debate," *American Banke*r, July 22, 1998; "Perils of Using Statistics to Show Presence or Absence of Loan Bias," *American Banker*, Jan 3, 1997; "Statistical Anomaly Penalizes Fair-Lending Effort," *American Banker,* Nov. 18, 1996; "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,* Jun 28, 1993; "Divining Difference," *Chance* 1994;7(4):38-9,48; "Comment on McLanahan, Sorensen, and Watson's 'Sex Differences in Poverty, 1950-1980,'" *Sign*s 1991;16(2):409-13; "The Perils of Provocative Statistics," *Public Interest* 1991;102: 3-14; "An Issue of Numbers," *National Law Journal*, Mar. 5, 1990; "The 'Feminization of Poverty' is Misunderstood," *Plain Dealer*, Nov 11, 1987 (reprinted in *Current* 1988;302(May) and *Annual Editions: Social Problems 1988/89.* Dushkin1988).

the two relative differences tend to changes in opposite directions as the prevalence of an outcome changes, can be illustrated with a wide range of publicly available data. Published income data, for example, show that reducing poverty will tend to increase relative differences in poverty rates while reducing relative differences in rates of avoiding poverty. National Health and Nutrition Survey (NHANES) data on systolic blood pressure show that generally reducing systolic blood pressure will tend to increase relative differences in hypertension rates while reducing relative difference in rates of avoiding hypertension. NHANES data on folate level show that generally improving folate levels will tend to increase relative differences in rates of low folate while reducing relative differences in rates of adequate folate. Life table data show that over the course of ages where most deaths occur, the lower the age the greater will tend to be relative differences in rates of dying before reaching the age and the smaller will tend to be relative differences in rates of surviving beyond it. Credit score data on mortgage loan applicants show that the lower the credit score the greater will tend to be relative differences in failing to achieve it while the smaller will tend to be relative differences in achieving it. See the Collected Illustrations sub-page of SR.

Normally distributed data on test scores, however, may most usefully illustrate the patterns. This is particularly so because lowering of cutoffs has long been considered a means of reducing the discriminatory impact of employment and other tests on which some demographic groups have lower average scores than others and because views about test cutoffs may in fact underlie beliefs of the Department of Justice that relaxing school discipline standards will reduce racial differences in discipline rates or that reducing the frequency of adverse lending outcomes will reduce racial differences in experiencing those outcomes. Thus, consider a situation where at a particular cutoff pass rates are 80% for an advantaged group (AG) and 63% for a disadvantaged group (DG). At this cutoff AG's pass rate is 27% higher than DG's pass rate. If the cutoff is lowered to the point where 95% of AG passes the test, assuming normal test score distributions, DG's pass rate would be about 87%. AG's pass rate would then be only 9.2% higher than DG's pass rate. It is because lowering cutoffs reduces relative differences in pass rates that lowering cutoffs is considered a means of reducing the discriminatory impact of tests.

But, whereas lowering cutoffs tends to reduce relative differences in pass rates, it tends to increase relative differences in failure rates. In the situation just described, DG's failure rate was initially 1.85 times AG's failure rate (37%/20%). With the lower cutoff, DG's failure rate would be 2.6 times AG's failure rate (13%/5%).

Figure 1 below, which like the numbers just cited is based on a situation where two groups have normally distributed test scores with means that differ by half a standard deviation, illustrates the effects on the two relative differences of serially lowering a cutoff from a point where almost everyone fails to a point where almost everyone passes. The numbers on the x-axis are the failure rates for the advantaged group, which are used as benchmarks for the overall prevalence of an outcome. The line with the diamond marker (blue in the electronic version of this letter) represents the ratio of DG's failure rate to AG's failure rate. The line with the square marker (red

in the electronic version) represents the ratio of AG's pass rate to DG's pass rate.[3] Thus, moving from the left to right, one observes that as test failure becomes less common, the relative difference in failure rates increases while the relative difference in pass rates decreases.

Figure 1. Ratios of (1) DG Fail Rate to AG Fail Rate and (2) AG Pass Rate to
DG Pass Rate at Various Cutoffs Defined by AG Fail Rate
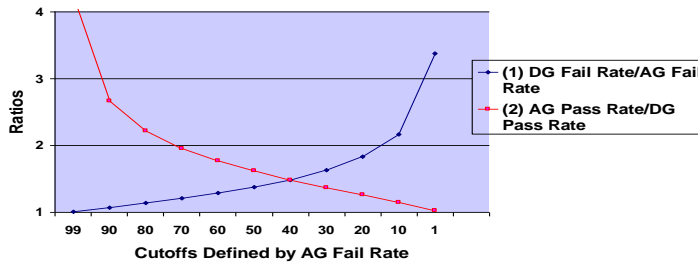
The features of the distributions underlying this pattern are illustrated in Figure 1 of the fall 1994 *Chance* article "Divining Difference."

School discipline standards operate just like test cutoffs. Less stringent standards, like lower test cutoffs, while leading to smaller relative differences in rates of avoiding discipline, lead to larger relative differences in discipline rates. Thus, discipline standards that are more lenient than existing ones will tend to increase the relative differences in discipline rates that are causing the concern. This subject is treated more fully, including discussion of some of the Department of Education's recently released data, on the Discipline Disparities page of jpscanlan.com. The Los Angeles SWPBS sub-page of the Discipline Disparities page discusses the situation where, after implementation of a program to reduce discipline rates, racial differences in discipline rates increased in South Los Angeles public schools. The Suburban Disparities sub-page discusses reportage of much higher racial differences in discipline rates in the suburbs of Philadelphia than in the city itself, explaining that comparatively large relative differences in adverse outcomes are common in settings where such outcomes are rarer than in other settings.

Lending standards also operate like test cutoffs. Standards that result in fewer adverse lending outcomes (including such things as rejection of a mortgage application and assignment to subprime rather than prime loan status), while leading to smaller relative differences in rates of

---

[3] It is common to use the disadvantaged group's rate as the numerator in the ratio used to derive relative differences in both pass rates and failure rates. I instead use the higher rate in both numerators, for several reasons including that the contrasting pattern of changes in the two relative differences is easier to recognize when both rate ratios are above (or below) one. While the choice of numerator will affect how one describes the size of a relative difference (e.g., when rates are 80% and 60%, the former is 33% greater than the latter while the latter is 25% less than the former), the choice of numerator is irrelevant to the points made here. Compare the illustration in Figure 1 of this letter with the illustrations based on income data in Figures 1 and 2 of the spring 2006 *Chance* article "Can We Actually Measure Health Disparities," which uses the disadvantaged group's rate as the numerator in both rate ratios.

securing favorable lending outcomes, lead to larger relative differences in adverse lending outcomes. This subject is treated more fully on the Lending Disparities page of jpscanlan.com. That page also addresses some particular aspects of the analysis underlying the Department's complaint in the *Countrywide* case.

I call your particular attention to the Lathern v. NationsBank sub-page of the Lending Disparities page. The *Lathern* case, which is the subject of the January 1, 1996 *Legal Times* article "When Statistics Lie," strikingly illustrates the anomaly in a situation where, after federal agencies urge lenders to adopt policies that those agencies should have recognized would tend to increase relative differences in rejection rates, lenders are targeted for litigation on the basis of the size of such differences. *Lathern* was a private action. I have not studied the extent to which the Department of Justice identified subjects for litigation on the basis of the size of relative differences in mortgage rejection rates or other adverse lending outcome. But few things are more inconsistent with responsible law enforcement than that the federal government should urge entities subject to federal law to adopt policies that make them more likely targets for litigation. Yet, at least since the March 8, 1994 issuance of the Policy Statement of the Interagency Task Force on Fair Lending, the federal government has been doing precisely that. As discussed in the recent *National Law Journal* article "The Lending Industry's Conundrum," the Department of Justice's *Countrywide* complaint continues this anomaly.

Similar issues, of course, exist in the case of schools and school districts that relax discipline policies based on the federal government's leading them to believe that doing so will reduce the relative differences in discipline rates on which schools and school districts will be judged both as to the impact of their of discipline policies and whether racial bias influences the administration of those policies.[4]

B. Other Statistical Issues Pertinent to the Enforcement of Civil Rights Laws

The failure to recognize the ways that measures of differences between outcome rates are affected by the overall prevalence of an outcome can undermine other aspects of the Department of Justice's efforts to enforce the nation's civil rights laws and ensure the fairness of the criminal justice system.

Section B of the Discipline Disparities page and the June 23, 1993 *Legal Times* article "Getting it Straight When Statistics Can Lie" address certain misperceptions about relative differences in adverse outcome rates in the employment setting. Section C of that page and the October 12, 1996 *Legal Times* article "Mired in Numbers" address certain misperceptions about relative differences in adverse outcome rates in the criminal justice system. Other pages on jpscanlan.com and other published articles listed in note 1 or made available jpscanlan.com also address statistical issues that may be pertinent to the Department of Justice's mission. At this point, however, I call your attention principally to such materials that are related to the issues

---

[4] A letter somewhat similar to this one addressing public school discipline policies and certain other matters of concern to the Department of Education was sent to that agency on April 18, 2012.

addressed in Section A above and that are relevant to the Department's mission with which I am most familiar, the enforcement of employment discrimination laws.[5]

A key point of the references in the third paragraph of Section A is that in order for a measure to effectively appraise the size of the difference in the well being of two groups reflected by a pair of outcome rates, the measures must be unaffected by the overall prevalence of an outcome. The Solutions sub-page of MHD describes such a measure, which involves deriving from a pair of outcome rates the difference between means of hypothesized normal underlying distributions of factors associated with experiencing the outcome. While I have generally discussed the approach in the context of the measurement of differences in health and healthcare outcomes, the principles apply as well in appraising the size of differences in employment outcomes.

The Relative Versus Absolute sub-page of MHD, which is directed at refuting a notion in health disparities literature that various measures of differences between outcome rates may each provide a valid appraisal of the size of a difference between pairs of rates even when they yield opposite conclusions as to the comparative size of differences in different settings, uses as an example a situation where the question to be answered involves which of several employers is the most biased. Table 1 below presents an abbreviated and slightly modified version of Table 1 of the Relative Versus Absolute sub-page.[6] The table below shows the hire and rejection rates of advantaged and disadvantaged groups at four employers, along with the rate ratios for hire and the rate ratios for rejection. In a situation where it assumed that for each employer the qualifications of the applicants from the advantaged group do not differ from the qualifications of the disadvantaged group and all differences in rates result from employer bias, the question posed is how might the employers be ranked according to level of bias.

---

[5] Issues raised in a body of work made available on the Employment Discrimination page of jpscanlan.com, including the issues raised in the fall 1988 *Public Interest* article "Illusions of Job Segregation" and the March 27, 1995 *National Law Journal* article "Multimillion-Dollar Settlements May Cause Employers to Avoid Hiring Women and Minorities for Less Desirable Jobs to Improve the Statistical Picture," are also relevant to the Department's enforcement of employment discrimination laws. As discussed in Section 8 of the Lending Disparities page, issues raised in the 1988 *Public Interest* article may also raise question about the analyses underlying the *Countrywide* complaint. But I am limiting this letter largely to issues involving the correlations between the prevalence of an outcome and measures of differences in rates of experiencing or avoiding it.

[6] Table 1 in the Relative Versus Absolute page presents relative differences – *i.e.*, rate ratio (RR) minus 1 for RRs greater than 1 and 1 minus RR for RRs less than 1 – and uses the disadvantaged group's rate as the numerator in both rate ratios. For consistency with Figure 1, Table 1 presents rate ratios and uses the higher number in the numerator for both rate ratios. See note 3 *supra*.

Table 1 in the Relative Versus Absolute sub-page also presents information on absolute differences between rates and odds ratios. As reflected on that sub-page (and as discussed with regard to the *Countrywide* complaint in Section 6 of the Lending Disparities page) issues about the way absolute differences and differences measured by odds ratios are affected by the prevalence of an outcome can be important, particularly with regard to questions such as whether pay-for-performance programs tend to increase or decrease healthcare disparities. See the Pay for Performance and Between Group Variance sub-page of MHD. But treatment of such issues is unnecessary in this letter.

Table 1. Illustrations of Associations of Measure of Differences between Outcome Rates with Prevalence of the Outcome

Employer	AGHireRate	AGRejectionRate	DGHireRate	DGRejectionRate	RRHire	RRRejct
A	20.00%	80.00%	9.00%	91.00%	2.22	1.14
B	40.10%	59.90%	22.70%	77.30%	1.77	1.29
C	59.90%	40.10%	40.50%	59.50%	1.48	1.48
D	90.00%	10.00%	78.20%	21.80%	1.15	2.18

The table shows that rankings of degree of bias based on relative differences in hire rates would be the exact opposite of rankings based on relative differences in rejection rates. Yet it obviously makes no sense to say one employer is more biased with respect to hiring applicants from the two groups and another is more biased with respect to rejecting applicants from the two groups. As is similarly shown in the Relative Versus Absolute subpage, there exists only one reality as to the comparative size of the difference in the circumstances of demographic groups reflected by two or more pairs of rates of experiencing an outcome. That is so with regard any type of outcome and regardless of the nature of the forces driving the difference in outcome rates. But it is most obviously so when the force may be bias against a demographic group.

According to the specifications underlying the table, which are the same as those underlying Figure 1, the degree of bias is exactly the same at each employer. That is, the best estimate of such bias in each case involves deriving from the pairs of hire or rejection rates the difference between means of the hypothesized underlying distributions, which in this case is half a standard deviation. The same reasoning would hold if the comparative qualifications of applicants from the two groups at each employer was unknown and the question posed concerned a ranking according to extent of differences in qualification necessary to explain the observed outcomes as other than a result of bias.

Many employment discrimination cases are based on comparisons of the proportion a group comprises of persons potentially experiencing an outcome (e.g., a labor market, job applicants, incumbent employees) and the proportion the group comprises of persons experiencing the outcome (e.g., hire, promotion, discipline, termination).[7] On the basis of such information, one can derive the relative rates of experiencing the outcome even though one does not know the actual rates of experiencing the outcome. For example, if a group comprises 40% of applicants and 20% of hires, it is possible to determine that such group's hire rate is 37.5% of that of persons not in the group (20/40 over 80/60), or, put another way, that persons not in the group are 2.67 as likely to be hired as persons in the group. But, as indicated in the immediately preceding paragraphs, it is necessary to know the actual underlying outcome rates in order to effectively appraise the size of differences in outcomes rates. See Table 1 of the 2009 Royal Statistical Society presentation for an illustration of the varying meanings of a particular ratio of rates of experiencing an outcome at different levels of overall prevalence of the outcome.

---

[7] Racial profiling analyses are invariably based on such comparisons. See the unpublished paper "The Profiling Conundrum."

This matter is illustrated more fully on the Representational Disparities sub-page of SR. See also the Gender Differences in DADT sub-page of the Vignettes page, which discusses the same issue in the context of reported perceptions concerning gender differences in the enforcement of the "don't-ask-don't-tell" policy in the military.[8]

The Employment Tests sub-page of SR explores whether, given the theories generally expressed on the Measuring Health Disparities and Scanlan's Rule pages, lowering a cutoff in fact reduces the disparate impact of a test in a meaningful way and explains why it does (assuming selection among persons who pass the test is not a function of test scores). I note, however, that the reasoning on that page would leave open whether, for example, in other contexts policies that generally yield lower adverse outcome rates than other policies should in fact be deemed to have a smaller adverse impact in the way that a test can be deemed to have a smaller adverse impact with a lower cutoff than with a higher cutoff.

C. Technical Issues

Since the Department might seek the views of statistical experts on the points made in this letter, I note that most questions or objections such experts might raise are addressed in references already provided, especially the Scanlan's Rule page its sub-pages. But I nevertheless add the following, somewhat technical points.

While I commonly describe the way standard measures of differences between outcome rates tend to be affected by the overall prevalence of an outcome as a consequence of the shape of normal distributions, 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 it will be present even when the distributions are not normal, as long as the distributions are not strongly irregular. For example, the pattern would be present when the distributions are uniform (i.e., rectangularly shaped). And Figures 8 and 10 of the 2008 International Conference on Health Policy Statistics presentation and Figure 1 referenced on the Credit Score Illustrations sub-page of SR illustrate that the patterns hold in the case of distributions known not to be normal because they are truncated parts of larger distributions. Thus, while the fact that the precise contours (and sometimes even the broad contours) of the underlying distributions may not be known will complicate efforts to appraise the size of differences in the well-being of two group reflected by two outcome rates in a way that is unaffected by the overall prevalence of an outcome, such fact by no means justifies relying on standard measures of differences between outcome rates as if those measures were unaffected by the overall prevalence of an outcome.[9] Such nuances as these aside, however, I trust that capable statistical experts, once thinking the

---

[8] Points similar to those made on the Relative Versus Absolute, Representational Disparities, and Gender Differences in DADT sub-pages are presented in the Case Study and Case Study Answers sub-pages of SR.

[9] See also the Illogical Premises and Illogical Premises II sub-pages of SR, which explains that, irrespective of distributional considerations, relative differences in either favorable or adverse outcomes are not merely unsound measures of association, but illogical ones. For when the relative difference for one outcome is the same in two settings involving different baseline rates, the relative difference for the other outcome necessarily will be different in the two settings.

matter through, will affirm that reducing the general prevalence of an outcome – whether involving employment, lending, school discipline, or anything else – will not tend to reduce relative differences in experiencing the outcome, but will tend to increase such differences

Some of the references collected in Section E.7 of MHD, which section addresses the extent of scholarly agreement with the views set out above, suggests that some statisticians may mistakenly focus on the many observed departures from the patterns I have described. But, as I think is made clear enough in the 2000 *Society* article "Race and Mortality" and many other places, one will of course observe departures from the described prevalence-related patterns because actual patterns will be functions of both (a) differences between the distributions in the settings being compared and (b) the prevalence-related forces. Is the understanding of (a) that is society's concern whether in the context of determining whether a disparity in health or healthcare outcomes has increased or decreased or in the context of determining which entity is most likely to have engaged in some form or proscribed discrimination. Only with a firm understanding of the prevalence-related forces implicated in observed patterns can one soundly address that concern.

Finally, some might make the related point that actions taken to reduce adverse outcomes rates can affect different groups in a meaningfully different way. That is certainly true. And, for example, actions that reduce any bias in the imposition of school discipline at the same time that they are reducing overall discipline rates will tend to reduce both relative differences in discipline rates and relative differences in rates of avoiding discipline (offsetting somewhat the prevalence-related patterns as to the former relative differences and enhancing somewhat the prevalence-related pattern as to the latter relative differences). But that is a different matter from the statistical points made here, which involve the forces that one must understand in order to distinguish changes that are functions of those forces from those that are not.

I hope that the Department will carefully consider the points made above and in the various references in carrying out its functions involving the enforcement of civil rights law or other matters involving the interpretation of data on demographic differences.

Best regards,

/s/James P. Scanlan

James P. Scanlan

cc:

Anurima Bhargava, Chief
Education Section
Civil Rights Division

Delora L. Kennebrew, Chief
Employment Section
Civil Rights Division

Steven H. Rosenbaum, Chief
Housing and Civil Enforcement Section
Civil Rights Division

[This letter is discussed on the [Holder/Perez Letter](#) sub-page of the [Lending Disparities](#) page of jpscanlan.com. Because the content of the letter may prove useful to reference in other materials, a link to this corrected copy is made available on the [Institutional Correspondence](#) sub-page of the [Measuring Health Disparities](#) page. The original is also available [here](#). The corrections are as follows:

[Page 5, second full paragraph, second last line: "as the" changed to "as to the"

[Page 8, fourth last line of body of letter: "group" changed to "groups"

[Page 9, first full paragraph, eighth line: "Is the" changed to "It is the"

[The corrections are sufficiently minor that a corrected copy was not mailed to the original recipients.]