

**UNITED STATES DISTRICT COURT
EASTERN DISTRICT OF MISSOURI
EASTERN DIVISION**

UNITED STATES OF AMERICA,)	NO. 4:16-cv-000180-CDP
)	
Plaintiff,)	JUDGE CATHERINE D. PERRY
)	
v.)	
)	
CITY OF FERGUSON,)	
)	
Defendant.)	
)	
)	
<hr style="width:40%; margin-left:0;"/>)	

**SUBMISSION OF JAMES P. SCANLAN
REGARDING PROPOSED CONSENT DECREE**

(Revised Apr. 15, 2015)

This is a shortened version of the document filed with the court on April 12, 2015. The original filing, which exceeded the page limit may, may be found [here](#). This version was sent to the court on April 15, 2016.

A central point of this submission, reflected well enough in both versions, involves the fact that the Department of Justice’s case against Ferguson, Missouri is based on the mistaken premise that unwarrantedly aggressive policing practices and unjustifiably harsh court practices caused African Americans to make up a higher proportion of persons experiencing adverse outcomes at the hands of the police and the courts than would be the case with less aggressive policing practices and less harsh court practices. In fact, the exact opposite is the case, as I explained in “Things [DoJ doesn’t know about racial disparities in Ferguson](#),” The Hill (Feb. 22, 2016). The government’s failure of understanding addressed in that article (and in numerous other places) is quite important to the fairness of the decree. Among other things, requirements in the decree will put the City in the situation of not knowing whether by reducing the frequency of an adverse outcome, and thereby increasing relative differences in rates of experiencing the outcome while reducing relative differences in rates of avoiding the outcome, it will be deemed to have increased or decreased the disparate impact of a practice with a disparate impact. An extended treatment of the principal statistical issues may be found in my October 8, 2016 [letter](#) to the American Statistical Association, urging the organization, among other things, to explain to arms of the United States Government that reducing the frequency of an adverse outcome tends to increase, not decrease, (a) relative differences between rates at which advantaged and disadvantaged groups experience the outcome and (b) the proportion disadvantaged groups make up of persons experiencing the outcome.

Table of Contents

	Page
Introduction and Summary	1
A. Patterns by Which the Two Relative Differences and Related Disparity Indicators Tend To Be Affected by the Frequency of an Outcome	3
B. Implications of the Patterns by Which Measures Tend to Be Affected by the Frequency of an Outcome With Respect to Particular Requirements of the Consent Decree	8
C. Suggested Disposition of this Matter	10
Table 1. Illustration of effects on relative differences in pass and fail rates, and proportions DG makes up of persons who pass and fail, of lowering a cutoff from a point where 80% of AG passes to a point where 95% of AG passes (when mean scores differ by approximately half a standard deviation and DG comprises 50% of test takers)	2
Figure 1. Ratios of (1) DG Fail Rate to AG Fail Rate and (2) AG Pass Rate to DG Pass Rate at Cut Points Defined by AG Fail Rates	4

Introduction and Summary

This submission¹ addresses the fact that actions of the Department of Justice (DOJ) Ferguson, Missouri have been based on an understanding of the relationship between the frequency of an outcome and the proportion African Americans make up of persons experiencing the outcome that is the exact opposite of reality. It also addresses the implications of the DOJ's misunderstanding of this relationship to the City of Ferguson's disparity monitoring obligations under the proposed Consent Decree.

For more than twenty years the DOJ and other agencies enforcing federal fair lending laws have encouraged lenders to relax lending standards and otherwise reduce the frequency of adverse borrower outcomes in order to reduce relative (percentage) racial/ethnic differences in rates of experiencing those outcomes. For at least several years, the DOJ and other federal agencies have encouraged public schools to relax discipline standards and otherwise reduce the frequency of adverse school discipline outcomes in order to reduce relative racial and other demographic differences in rates of experiencing those outcomes.

The belief that reducing the frequency of such outcomes will reduce relative differences in experiencing them, however, is the exact opposite of reality. Reducing the frequency of an adverse outcome will tend to reduce relative differences in rates of experiencing the corresponding favorable outcome. But reducing the frequency of an adverse outcome will tend to increase relative differences in rates of experiencing the outcome itself. Unaware that reducing the frequency of an outcome tends to increase relative differences in rates of experiencing it – indeed, believing just the opposite – the government has continued to monitor the fairness of lending and school discipline practices on the basis of relative differences in adverse outcomes. Thus, lenders and schools that reduce the frequency of adverse outcomes in compliance with government encouragements increase the chances that the government will accuse them of discrimination.

Reducing the frequency of an outcome also tends to increase the proportions groups most susceptible to the outcome make up of persons (a) experiencing the outcome and (b) failing to experience the outcome. Here, too, as reflected by the DOJ's actions regarding Ferguson, Missouri, the government has consistently acted on the mistaken belief that reducing the frequency of an adverse outcome will tend to reduce the proportions groups most susceptible to the outcome make up of persons experiencing the outcome. Thus, here, too, the government has been encouraging entities covered by civil rights law to take actions that increase the chances that the government will accuse them of discrimination.

A succinct treatment of these issues with a particular focus on the instant litigation may be found in my "[Things DoJ doesn't know about racial disparities in Ferguson](#)," *The Hill* (Feb. 22, 2016).

¹ This submission is a revision of an over-length item filed on April 12, 2016 ([Original Submission](#)). Underlinings in print copies of this letter, as in the case of the item just mentioned, reflect links to referenced materials in electronic copies of the letters that may be found by means of the [Institutional Correspondence](#) subpage of the [Measuring Health Disparities](#) page of [jpscanlan.com](#).

The fact that reducing the frequency of an outcome tends to increase, not decrease, relative difference in rates of experiencing the outcome and the proportions more susceptible groups make up of persons experiencing the experiencing, while not widely known, is hardly debatable. Confronted with this submission and required to address the points it makes, the DOJ should acknowledge that it has been mistaken on the matter.

But whether the DOJ acknowledges its mistaken understanding or not, the Consent Decree in its current form – with its various provisions requiring the City of Ferguson to monitor racial differences and its various provisions aimed at generally reducing adverse outcomes, all in a context where the government itself does not understand the statistical consequences of reducing the frequency of an outcome – would place the City in an untenable situation. Thus, approval should be deferred until various statistical issues are addressed and resolved.

Section A of this submission briefly describes the above-mentioned statistical issues, while referencing materials that address the matters in greater detail. Section B addresses certain statistical issues having particular bearing on the City of Ferguson’s obligations under the decree. Section C suggests that the Court defer approval of the decree until the DOJ addresses the measurement issues discussed in Sections A and B.

A. Patterns by Which the Two Relative Differences and Related Measures of Disparity Tend To Be Affected by the Frequency of an Outcome

There are four principal measures by which observers appraise differences in 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 (3) absolute (percentage point) differences between outcome rates; and (4) odds ratios. None of these measures provides a sound basis for quantifying differences in the circumstances of advantaged and disadvantaged groups reflected by their outcome rates (or, otherwise put, the strength of the forces causing the groups’ outcome rates to differ) because, for reasons related to features of underlying risk distributions each measure tends to be systematically affected by the frequency of an outcome.

The above measures most pertinent to disparate impact and other discrimination issues are the two relative differences and I will limit the discussion here to those measures. The pertinent statistical pattern with respect to the two relative differences is that whereby the rarer an outcome the greater tends to be the relative difference in experiencing it and the smaller tends to be the relative differences in avoiding it.²

In some settings, including the instant suit, appraisal of demographic differences for purposes of analyzing discrimination or other issues are based, not on a comparison of outcome rates for different groups, but on a comparison of the proportion a group makes up of persons potentially experiencing an outcome and the proportion the group makes up of persons actually experiencing the outcome. Such comparisons are affected by the frequency of an outcome in a

² 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.”

manner related to the above-described pattern regarding relative differences. Specifically, 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 is a pattern whereby the rarer an outcome the greater tend to be the proportions groups most susceptible to the outcome make up of both persons experiencing the outcome and persons avoiding the outcome.

These patterns can be easily illustrated with normally distributed test score data. Table 1 below, 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 at two cutoff points, 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).³ 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, and proportions DG makes up of persons who pass and fail, of lowering a cutoff from a point where 80% of AG passes to a point where 95% of AG passes (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. Thus, the ratio of AG’s pass rate to DG’s pass rate would be 1.27. 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%. Thus, lowering the cutoff reduces the ratio of AG’s pass rate to DG’s pass rate from 1.27 to 1.08.

The fact that lowering a test cutoff tends to reduce relative differences in pass rates is widely known and underlies the common understanding that lowering a test cutoff reduces the disparate impact of employment and other tests where some demographic groups outperform others.⁴ Such fact may also underlie the government’s longstanding (though mistaken) belief that relaxing standards will tend to reduce relative differences in adverse outcomes.

³ While I commonly refer to patterns of relative differences in this submission, the tables I use actually present rate ratios. Where the larger figure is used as the numerator in the rate ratio, the relative difference is the rate ratio minus 1.

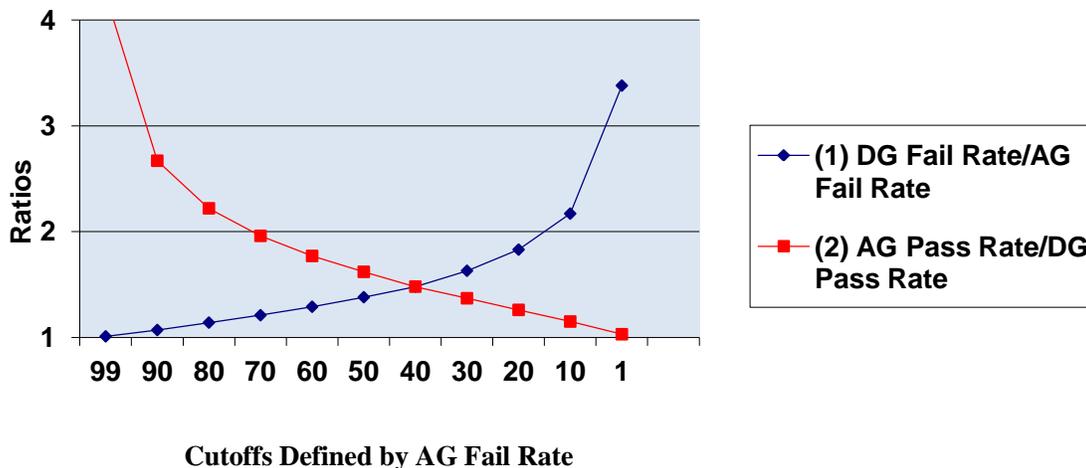
⁴ Whether lowering (or raising) a standard in fact increases or decreases the impact of a standard is a complex issue and the answers may vary from setting to setting. See Section E (at 27-32) of the Kansas Law [paper](#) discussed *infra*.

But, while lowering a cutoff tends to reduce relative differences in pass rates, it tends to increase relative differences in failure rates. With the higher cutoff, DG's failure rate was 1.85 times AG's failure rate (37%/20%). Lowering the cutoff increases that ratio to 2.60 (13%/5%).

The final two columns of the table show how lowering the cutoff increases the proportions DG makes up of persons passing the test and persons failing the test. Assuming, DG makes up half of the test takers, the former proportion would increase from 44% to 48% and the latter proportion would increase from 65% to 72%.

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 is based on 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 and (2) AG Pass Rate to DG Pass Rate at Cutoff Points Defined by AG Fail Rates



A graphical illustration of the pattern of changes in the proportion DG makes up of person who pass and persons who fail across the full range of cutoffs may be found in Figure 1a (at 18) of the applied statistics [workshop](#) at Harvard's Institute for Quantitative Social Science identified in note 5 *infra*.

The patterns illustrated above, however, are not limited to test score data or the numbers I chose to illustrate them. Rather, they are found in virtually all situations where groups differ in their susceptibility to an outcome.

The statistical patterns discussed in this submission are addressed at much greater length in an October 8, 2015 letter to the [American Statistical Association](#) that, among things, urges the organization to explain to arms of the United States Government that their beliefs about the statistical implications of reducing the frequency of an outcome are mistaken. Other recent

extended treatments of the issues 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); and the TDHCA [brief](#) mentioned in the Statement of Interest. See also my forthcoming “The Mismeasure of Health Disparities,” *Journal of Public Health Management and Practice* (July/Aug. 2016) (available online in May 2016).

Extensive graphical and tabular illustrations of the pertinent patterns may be found in methods workshops given at arms of American universities between 2012 and 2015,⁵ as well as in more than a score of presentations given at statistics, epidemiology, demography, and public health conferences in North America and Europe between 2001 and 2011 (available [here](#)). Treatments of the patterns with respect to particular outcomes, usually with tabular illustrations, may also be found on the pages and subpages of [jpscanlan.com](#) devoted to measurement issues.⁶ Over 140 online comments (available [here](#)) address the problems in various studies in medical, epidemiological, and health policy journals that attempted to measure aspects of health or healthcare disparities – or to provide guidance on the interpretation of subgroup effects or the calculation of number-needed-to-treat – without consideration of the patterns by which measures tend to be affected by the frequency of an outcome.

Recent, fairly succinct treatments of the issues in the context of the government’s encouraging lenders and public schools to relax lending and discipline standards under the mistaken belief that doing so will tend to reduce relative racial differences in adverse borrower and discipline outcomes, as well as the DOJ’s actions regarding Ferguson, Missouri, may be found in the March 2016 *Hill* article mentioned above, as well as [“Things DOJ doesn’t know about racial disparities in Ferguson,”](#) *The Hill* (Feb. 22, 2016) (discussed above); [“Things government doesn’t know about racial disparities,”](#) *The 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). Older, somewhat longer treatments of

⁵ See [“The Mismeasure of Health Disparities in Massachusetts and Less Affluent Places,”](#) Department of Quantitative Health Sciences, University of Massachusetts Medical School (Nov. 18, 2015) [“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).

⁶ The principal measurement pages are: [Measuring Health Disparities](#), [Scanlan’s Rule](#), [Mortality and Survival](#), [Statistical Reasoning](#), [Immunization Disparities](#), [Immunization Disparities](#), [Educational Disparities](#), [Disparate Impact](#), [Discipline Disparities](#), [Lending Disparities](#), [Employment Discrimination](#), [Feminization of Poverty](#). The pages have close to 100 subpages.

the misunderstanding of the effects of relaxing standards on measures of disproportionality, which involve a variety of settings and illustrate both how longstanding and how universal is that misunderstanding, may be found in my "[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); and "[An Issue of Numbers](#)," *National Law Journal* (Mar. 5, 1990).

Recent, fairly succinct treatments of the misunderstanding of the implications of reducing the frequency of an adverse outcome on relative differences in experiencing the outcome that discuss the potential unconstitutionality of disparate impact claims under the Fair Housing Act may be found in my "[Is HUD's Disparate Impact Rule Unconstitutionally Vague?](#)," *American Banker* (Nov. 10, 2014) and "[Case may reveal government's perverse fair lending enforcement](#)," *The Hill* (Dec. 29, 2014), an issue also treated in the TDHCA brief (at 31-32). A key point in the analysis of that issue involves the fact that the Department of Housing and Urban Development's regulations require that, even when practices with a disparate impact can be justified by sound business purposes, the covered entity must adopt the least discriminatory alternative that equally serves those purposes (which commonly, in the view of the government and others, involves relaxing a standard). But covered entities are provided no guidance on whether appraisals of the comparative size of a disparate impact will be based on the relative difference in the favorable outcome or the relative differences in the adverse outcome (though probably they will be judged on the measure that relaxing a standard tends to increase).

The same points would apply to any disparate claim in circumstances where it is unknown whether reducing the frequency of an outcome will be regarded as reducing the disparate impact of the outcome on the basis of the relative difference in rates of avoiding the outcome or as increasing the disparate impact on the basis of the relative difference in experiencing the outcome. The Civil Rights Act of 1991 specifically requires that employers adopt the practices with the least disparate impact, which in the employment context is sometimes measured in terms of the relative difference in the favorable outcome and sometimes measured in terms of the relative difference in the adverse outcome. There was no Congressional intent with respect to which relative difference should be used to measure impact because then, as now, Congress was unaware that it was even possible for the two relative differences to change in opposite directions as a standard is altered, much less that this tends to occur systematically.

A particularly problematic feature of the proposed Consent Decree in this case may be found in its Paragraph 74 (at 18-19), which specifically requires that the Ferguson Police Department identify and implement any available alternative to a "program, initiative, activity, or service" having a disparate impact that has "less of a disparate impact." All other issues aside, no Consent Decree should be entered unless it specifies how it will be determined whether one practice has a less of a disparate impact than another.

The Consent Decree's Paragraph 285 (at 65) also requires that the Ferguson Police Department employ hiring practices with the least disparate impact. It is unclear whether the requirement imposes any obligation beyond that imposed by Title VII of the Civil Rights Act.

But if such a provision is to be incorporated into a decree, the decree should clarify the measurement issue.

Treatments of the above-described and related patterns may also be found in letters to governmental and nongovernmental entities addressing the ways that analyses of demographic or other differences in outcome rates that the entities conduct, fund, or provide guidance on – or that in some manner pertain to the entities’ activities – are undermined by the failure to recognize patterns by which measures of differences between outcome rates tend to be affected by the frequency of an outcome. In addition to illustrating the patterns (and failures to understand them) in particular settings and with regard to particular activities, the letters reflect the pervasiveness of the misunderstanding of the patterns even among the nation’s most prestigious institutions.⁷

The pattern, however, was recognized by statisticians of the National Center for Health Statistics as early as 2004.⁸ Scholars outside the United States Government have recognized the pattern since 2005.⁹ As suggested above, while few people are familiar with the patterns described above, the existence of the patterns is not open to dispute.

⁷ Such letters include the following (some of which have already been mentioned): [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), [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)

⁸ (a) Keppel KG, Percy JN, Klein RJ. [Measuring progress in Healthy People 2010](#). Healthy People Statistical Notes. No. 25. Hyattsville, MD: National Center for Health Statistics. 2004. (b) Keppel KG, Percy JN. Measuring relative disparities in terms of adverse events. *J Public Health Manag Pract* 2005;11(6):479–483; (c) Keppel KG, Pamuk E, Lynch J, et al. [Methodological issues in measuring health disparities](#). *Vital Health Stat* 2005;2 (141); (d) Keppel KG, Percy JN. Healthy People 2010: Measuring Disparities in Health. *Chance* 2009;22(1):6-7. See also recent recognition that relative differences in See also recent recognition that relative differences in experiencing an outcome tend to be large where the outcome is uncommon by scientists from the Centers for Disease Control and Prevention (apart from NCHS) in Moonesinghe R, Beckles GLA. [Measuring health disparities: a comparison of absolute and relative disparities](#). *PeerJ*. 2015;3:e1438; DOI 10.7717/peerj.1438.

⁹ (a) Carr-Hill R, Chalmers-Dixon P. [The Public Health Observatory Handbook of Health Inequalities Measurement](#). Oxford: SEPHO; 2005; (b) Houweling TAJ, Kunst AE, Huisman M, Mackenbach JP. [Using relative and absolute measures for monitoring health inequalities: experiences from cross-national analyses on maternal and child health](#). *International Journal for Equity in Health* 2007;6:15; (c) [Eikemo TA, Skalicka V, Avendano M. Variations in health inequalities: are they a mathematical artifact?](#) *International Journal for Equity in Health*

This does not mean that there will be no departures from the patterns, as I have explained in numerous places. Observed patterns are functions of (a) the described frequency-related patterns and (b) the strength of the forces causing the outcome rates to differ in the settings being compared. The strength of those forces may vary greatly from setting to setting and may change substantially over time, particularly during periods of dramatic changes in the frequency of an outcome. Such factors may cause the frequency-related patterns not to be observed even though they are having a substantial effect. In fact, there are aspects of the proposed Consent Decree of a type that might tend to reduce all measures of differences between outcome rates and disproportionality, and do so sufficiently to that overall relative difference in certain adverse outcomes, and the proportion African Americans make up of persons experiencing those outcome to decrease. See the [Original Submission](#) at 22.¹⁰ The crucial consideration is that so long as measures tend to be affected by the frequency of an outcome in any manner, one cannot rely on the measures to appraise the strength of the forces causing rates to differ without consideration of the role of the frequency of the outcome.

In light of the above, and irrespective of the way the points made above or in the material cited above bear on perceptions about disparities in Ferguson warranting remedy, I suggest that it would be manifestly unjust to impose on the City the many disparity monitoring obligations in the Consent Decree without addressing the Department of Justice's misunderstanding of pertinent statistical issue.

B. Implications of the Patterns by Which Measures Tend to Be Affected by the Frequency of an Outcome With Respect to Particular Requirements of the Consent Decree

Assuming a decree were to be entered with the same disparity monitoring provisions of the proposed Consent Decree, it would be essential that those attempting to monitor those disparities, and those overseeing the monitoring, understand certain things about measuring disparities in a rational manner.

Most of the above-mentioned recent extended treatments of the patterns by which measures tend to be affected by the frequency of an outcome discuss a means of quantifying the difference between the circumstances of two group reflected by their favorable or adverse outcome rates. By way of example, the favorable or adverse outcome rates in each row of Table

2009;8:32; (d) Lambert PJ, Subramanian S. Disparities in Socio-Economic outcomes: Some positive propositions and their normative implications. *Soc Choice Welf* 2014;43:565-576); and (e) Lambert PJ, Subramanian S. [Group inequalities and "Scanlan's Rule": Two apparent conundrums and how we might address them](#). Working Paper 84/2014, Madras School of Economics (2014).

¹⁰ Nevertheless, the frequency-related patterns described here are substantial enough that they commonly will be observed even if the strength of the forces causing outcome rates to differ changes substantially. This evidenced by the fact that, even though school district that are generally reducing discipline rates are simultaneously employing varied measures to reduce racial differences in discipline rates, have typically shown increases in relative differences in discipline rates. See the following subpages of the [Discipline Disparities](#) page of [jpscanlan.com](#): [California Disparities](#), [Colorado Disparities](#), [Connecticut Disparities](#), [Maryland Disparities](#), [Minnesota 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](#)

1 would yield a difference between the underlying means of .5 standard deviations. Other useful illustrations may be found in the tables of "[Race and Mortality Revisited](#)" and the FCSM [paper](#) and the methods workshops listed in note 5 *supra*. Particularly instructive is Table 8 (at 342) of "Race and Mortality Revisited," which involved what the Departments of Justice and Education in March 2014 regarded as especially large racial disparities in preschool suspension rates without consideration of the role of the fact that suspensions are extremely rare in preschool.

But in order to employ this measure (or any other plausible measure) one must have the actual outcomes rates. Information solely on the proportion a group makes up of persons potentially experiencing the outcome and the proportion it makes up of persons actually experiencing the outcome, while enabling one to derive the relative difference between rates of experiencing that outcome, does not enable one to derive the rates themselves. See discussion of this issue in the Section B (at 23-26) of the Kansas Law [paper](#) and Section I.B (at 23-27) of the TDHCA [brief](#).

There are additional problems with appraisals of the differences in the circumstances of advantaged and disadvantaged groups based on comparisons of such proportions. These are touched up in the TDHCA brief (at 26-27) and addressed at length in the [IDEA Data Center Disproportionality Guide](#) and the [Ferguson, Missouri Arrest Disparities](#) subpages of the [Discipline Disparities](#) page of [jpscanlan.com](#). See also [slides 98-108](#) of the October 2014 University of Maryland [methods workshop](#) referenced in note 5 *supra*.¹¹ But since comparisons of the referenced proportions do not provide a sound method for appraising differences in the circumstances of two groups in any event (just as relative differences in either outcome do not¹²), it is unnecessary to belabor here the additional problems with such analyses.

One must be mindful, however, that in order to derive outcome rates for groups being compared, it is necessary to identify the appropriate numerators and denominators. In some

¹¹ One of those additional problems is that for any given pair of rates, the larger the proportion the subject group makes up of persons potentially experiencing the outcome, the smaller will be the relative difference between the proportion the group makes up of persons experiencing the outcome the proportion it makes up of persons actually experiencing the outcome (or the corresponding ratio of the former proportion to the latter proportion, often called the Disparity Index, as in the submission of commenter Keith Kallstrom). See Table 2 of [IDEA Data Center Disproportionality Guide](#) subpage and Table 24 (slide 108) of the University of Maryland [workshop](#). An illustration of this problem may be found in discussion in the Kallstrom submission (at 8) regarding the fact that the Disparity Index for Ferguson was only 1.37, which was much smaller than the average for the randomly selected jurisdictions whose data were studied. Given that the African American proportion of Ferguson's population is 67%, even if African Americans were subjects of 100% of traffic stops, the Disparity Index could not exceed 1.49. The largest Disparity Index shown in the table on page 2 of Mr. Kallstrom's submission is the 10.72 figure for Glendale, while the next largest is 5.02. The size of the Glendale Disparity Index is no doubt in some part a function of the fact that African Americans make up less than 1% of the population of Glendale.

¹² See "Race and Mortality Revisited" (at 339-41) and the American Statistical Association letter (at 12-13) regarding the fact that, irrespective of the patterns I describe above, the rate ratio (or its associated relative difference) is an illogical measure of association., See also the [Illogical Premises](#), [Illogical Premises II](#), [Subgroup Effects](#), [Subgroup Effects – Nonclinical](#), and [Inevitability of Interaction](#) subpages of the [Scanlan's Rule](#) page of [jpscanlan.com](#) and the [Comment on Hingorani BMJ 2013](#). For an illustration of this point with regard to the "four-fifths rule" of the Uniform Guidelines on Employee Selection Procedures, see the [Four-Fifths Rule](#) subpage of the [Disparate Impact](#) page of [jpscanlan.com](#)

cases, identification of those figures is straightforward. But it is not straightforward in the case of arrest data. See the Addendum to the above-mentioned [Ferguson, Missouri Arrest Disparities](#) subpage of the Discipline Disparities page of [jpscanlan.com](#). As reflected in the Addendum, I have yet to satisfactorily resolve this issue.¹³ But, as difficult as it may be to resolve this issue in order to employ a sound measure, such difficulty does not provide a reasonable basis for reliance on a fundamentally unsound measure.

See also the [Original Submission](#) at 24-34 regarding certain related issues.

C. Suggested Disposition of This Matter

In light of the points made above, including those pertaining to the key premise of the Department of Justice's actions regarding the City of Ferguson, I respectfully suggest that the Court defer decision on the approval of the Consent Decree until the Department of Justice (a) addresses whether it is mistaken in understanding of the statistical implications of reducing the frequency of adverse outcome and provides guidance on the monitoring of the decree and (b) specifies methods by which the City of Ferguson will monitor racial differences in outcomes (including with respect to whether one practice has less of a disparate impact than another), and by which the Department will oversee such monitoring, that take into account, or are unaffected by, the frequency of the outcome examined.

Date: April 15, 2016

Respectfully submitted,

/s/ James P. Scanlan

James P. Scanlan

¹³ See the Kansas Law paper (at 21) regarding a similar interpretive problem in the employment context. See also my "[Measuring Hiring Discrimination](#)," *Labor Law Journal* (July 1993), regarding this problem in the employment tester context.