

Is the Disparate Impact Doctrine Unconstitutionally Vague?

James P. Scanlan
The Federalist Society Blog (fedsoc.org/blog)
May 5, 2016

Due process requires that a statute apprise a person of reasonable intelligence of the nature of prohibited conduct. Otherwise the statute is unconstitutionally vague. The void for vagueness doctrine, while principally applied to criminal statutes and those affecting speech, is not limited to those contexts. But even if application to other contexts is limited to extraordinary situations, the doctrine may have a role with respect to statutes imposing liability for a practice's disparate impact. For, in that context, there exist situations where statutes do not apprise person of far above average intelligence of the nature of prohibited conduct and do not even apprise government agencies enforcing the statutes of the nature of the prohibited conduct. Remarkable consequences of such situations include longstanding patterns where the government encourages entities covered by civil rights laws to engage in conduct that increases the chances that the government will sue them for discrimination.

Under the disparate impact doctrine, even when not intended to discriminate on a proscribed basis, a practice that disproportionately disadvantages groups protected by civil rights statutes is unlawful unless it can be shown to have a sound justification. Though explicitly codified only in the Civil Rights Act of 1991, the doctrine has long been applied by the courts to a number of laws prohibiting discrimination on the basis of race and various other characteristics.

An essential element of the doctrine, according to statute, case law, and federal regulation, is that even a justified practice is unlawful if there exists an alternative practice that equally serves the user's legitimate interests with less of a disparate impact. The determination of whether one practice has less of a disparate impact than another, however, is rather more complicated than most have imagined. For methods deemed entirely legitimate by the legal and scientific communities will commonly yield opposite conclusions about whether modification of a practice increases or decreases its disparate impact.

Percentage Differences in Favorable Outcome Rates

The Uniform Guidelines on Employee Selection Procedures provide a rule of thumb, commonly termed the "four-fifth" or "80 percent" rule, whereby a disparate impact will be found where the disadvantaged group's favorable outcome rate is less than 80% of the advantaged group's favorable outcome rate. Consistent with an understanding that stringent criteria have greater disparate impacts on disadvantaged groups than more lenient ones, the Guidelines also make clear that an employer will have to ensure that employment test cutoffs are no higher than necessary for the job in question.

Consider, then, an employment test where the pass rates are 80% for whites and 63% for a racial minority. The minority pass rate is only 79% of the white pass rate. So the test would be deemed to have a disparate impact under the four-fifths rule.

Suppose that the cutoff is lowered to a point where 95% of whites pass the test. Assuming normal test score distributions, the minority pass rate would be about 87%. With the lower cutoff, the minority pass rate would be 92% of the white rate, and the test would seem no longer to violate the four-fifths rule.

The fact that lowering a test cutoff tends to reduce percentage differences in pass rates is why lowering cutoffs has long been considered a means of reducing the disparate impact of tests on which some groups outperform others. Such fact also presumably plays an important role in the aforementioned understanding that more stringent criteria have greater disparate impacts on disadvantaged groups than more lenient ones.

Percentage Differences in Adverse Outcome Rates

But there is also an aspect of the matter that virtually no one understands. One remarkable manifestation of the misunderstanding of this aspect may be found in the enforcement of fair lending laws. Since the early 1990s federal agencies enforcing such laws have been concerned that minorities tend to have their mortgage applications rejected several times as often as whites. The agencies have also recognized the role of traditional lending criteria in causing these differences. Thus, since at least 1994, those agencies laws have been encouraging lenders to relax lending criteria because of the potential for the criteria to disproportionately disadvantage minorities.

Relaxing lending criteria has the same effect as lowering test cutoffs and will tend to reduce percentage differences in rates at which whites and minorities have their mortgage applications approved. Thus, lenders who relaxed criteria in accordance with federal encouragements generally reduced those differences.

But let us consider the matter from a different perspective and one that in fact is more pertinent to the government's enforcement of fair lending laws. In the test score hypothetical, with the original cutoff, the failure rates were 20% for whites and 37% for minorities. The minority failure rate was thus 1.85 times (85% greater than) the white failure rate. With the lower cutoff, the failure rates are 5% for whites and 13% for minorities. So with the lower cutoff the minority failure rate is 2.6 times (160% greater than) the white failure rate.

Thus, while lowering the cutoff reduced the percentage difference in pass rates, it increased the percentage difference in failure rates. This pattern is not peculiar to test score data or the numbers I chose to illustrate it. Inherent in normal (or, indeed, all but highly irregular) distributions of factors associated with experiencing an outcome is a pattern whereby the rarer an outcome, (a) the greater tends to be the percentage difference in experiencing it and (b) the smaller tends to be the percentage difference in avoiding it (*i.e.*, experiencing the opposite outcome). The pattern can be illustrated with almost any data showing points on a continuum of factors associated with experiencing an outcome. Income and credit score data, for example, show that the lower an income or credit score requirement, the greater will be the percentage difference between rates at which advantaged and disadvantaged groups fail to meet the requirement and the smaller will be the percentage difference between rates at which such groups meet the requirement. The pattern is also reflected in the countless situations where reductions

in the frequency of an outcome have in fact been accompanied by increased percentage differences in rates of experiencing the outcome and decreased percentage differences in rates of avoiding the outcome.

There will be departures from the pattern whereby percentage differences in experiencing an outcome and avoiding the outcome change in opposite direction as the frequency of an outcome changes. For other factors are commonly at work. And, of course, an alternative practice such as a test where the minority pass rate is 70% when the white pass rate is 80% would show smaller percentage differences in both pass and fail rates at any cutoff point defined by the white pass rate than the test described above.

But typically one will observe the referenced pattern whenever there occurs a general change in the frequency of an outcome akin to that effected by the lowering of a test cutoff. And in circumstances where percentage differences in favorable outcomes are very small, percentage differences in adverse outcomes will tend to very large.

Thus, relaxing lending criteria, or doing anything that tends generally to increase rates of mortgage approval and other favorable borrower outcomes, while tending to reduce percentage differences in favorable outcome rates, will tend to increase percentage differences in the corresponding adverse outcome rates.

Federal agencies enforcing fair lending laws, however, have been utterly unaware that relaxing lending criteria tends to increase percentage differences in adverse borrower outcomes. In fact, they have acted under the belief that it will do just the opposite. And they have consistently monitored the fairness of lender practices on the basis of percentage difference in adverse borrower outcomes. Thus, in consequence of the government's failure to understand a fundamental statistical pattern, lenders that comply with government encouragements to relax standards tend to increase the chances that the government will sue them for discrimination.

Nearly Universal Failure of Understanding

The government is not alone in its failure to understand this matter. I have explained that reducing the frequency of an outcome tends to increase percentage differences in rates of experiencing the outcome, while reducing percentage differences in rates of experiencing the corresponding opposite outcome, in quite a few places since 1987. Statisticians for the National Center for Health Statistics recognized the pattern for the first time in 2004. Since then a handful of scholars here and abroad have also recognized it.

But the pattern continues to remain unknown among the vast majority of persons and institutions analyzing group differences in the law and the social and medical sciences. Indeed, as reflected by the scores or hundreds of articles observing that "despite" a general decline in mortality or some other adverse outcome, percentage demographic differences either "persist" or "have increased," the government's mistaken notion that reducing the frequency of an outcome will or should reduce percentage differences in rates of experiencing it is shared by a substantial part of the scientific community. As a result of the general failure to understand the way the two percentage differences (and certain other measures) in fact tend to be affected by the frequency

of an outcome, little that has been said about such things as whether some racial or other demographic disparity has increased or decreased over time, or even about whether it should be deemed large or small, has had a sound statistical basis.

The scientific community's failure to understand these issues has caused an immense waste of resources even when misinterpretations of data have not adversely influenced policy decisions. But it is the failure on the part of the federal civil rights establishment that has created the anomaly whereby an entity's following the government's guidance increases the chances that the government will sue the entity for discrimination.

In October 2015, I [asked](#) the American Statistical Association (ASA) to explain this issue to the various government entities whose policies reflect the mistaken belief that reducing the frequency of an outcome will tend to reduce percentage differences in rates of experiencing the outcome. It will take some time for ASA to understand the matter itself and even more time to decide to say anything about it. The same holds for other entities of seeming statistical expertise to which I have made similar requests, but whose leaderships have yet to understand that it is even possible for the two percentage differences to change in opposite directions as the frequency of an outcome changes, much less that they tend to do so systematically.

Let us assume, however, that the government will eventually recognize that relaxing standards tends to increase percentage differences in rates at which advantaged and disadvantaged groups fail to meet the standards. That could come easily enough in a court case if the defendant simply understood the matter well enough to pose appropriate interrogatories to the government plaintiff.

Causing the government to understand the pertinent statistical principles ought to eliminate the law enforcement anomaly, or so one would hope. But there would remain the question of whether relaxing a standard – and thereby reducing percentage differences in meeting the standard while increasing percentage differences in failing to meet the standard – reduces or increases the standard's disparate impact.

Varied Disparate Impact Anomalies

Before suggesting an answer to that question, I briefly describe some of the other areas where federal government policies are based on an understanding of statistics that is the opposite of reality. For several years the Departments of Justice (DOJ) Education (DOE), while referencing the disparate impact of stringent discipline policies, have encouraged public schools to relax discipline standards in order to reduce percentage racial/ethnic and other differences in adverse discipline actions like suspension and expulsion. But here, too, while reducing standards tends to reduce percentage differences in rates of avoiding adverse discipline actions, it tends to increase percentage differences in discipline rates. And, in fact, all across the country school systems that have been generally reducing discipline rates in response to government guidance, sometimes enacting laws to do so, have been finding increased percentage racial/ethnic differences in discipline rates.

Congress shares the mistaken view that reducing the frequency of an outcome tends to reduce percentage differences in experiencing the outcome. The Individuals with Disabilities Education

Act specifically provides that when there exist “significant discrepancies” in the suspensions of students with disabilities (typically measured in terms of percentage differences in suspension rates), school districts must consider implementing measures of the type that generally reduce suspension rates. The Keep Kids in School Act introduced in the Senate in 2015, and still in committee, reflects a like misunderstanding of the effects of reducing the frequency of suspensions on percentage differences in suspension rates.

Sometimes the government and others measure racial and other disparities (whether or not characterized as disparate impact), not in terms of differences between outcome rates, but in terms of the difference between (a) the proportion a group makes up of persons potentially experiencing an outcome and (b) the proportion it makes up of persons actually experiencing the outcome. One recent example may be found in the attention DOJ and DOE gave to racial differences in preschool suspensions in March 2014, while prominently citing the fact that blacks made up 48% of preschool students suspended more than once even though they were only 18% of preschool students. The same figures were emphasized in a December 2014 Policy Statement issued jointly by the DOE and the Department Health and Human Service (HHS) recommending a variety of measures aimed at (in the views of the two agencies) reducing the proportion blacks and other disadvantaged groups made up of suspended students by generally reducing preschool suspensions.

But a corollary to the whereby the rarer an outcome, the greater tends to be the percentage difference in experiencing it, and the smaller tends to be the percentage difference in avoiding it, is a pattern whereby the rarer an outcome, the greater tends to be the proportions groups most susceptible to the outcome make up of (a) persons experiencing the outcome and (b) persons avoiding the outcome. For example, in the test score hypothetical, assuming that minorities make up 50% of the test takers, lowering the cutoff in the manner described would (a) increase the proportion minorities make up of person failing the test from 65% to 72% and (b) increase the proportion minorities make up of persons passing the test from 44% to 48%. The pattern of directions of changes in the proportions minorities make up of persons failing and passing the test holds regardless of the proportion minorities make up of test takers.

Thus, while the DOE/HHS Policy Statement asserts that preschool suspension rates are “high,” it is the fact that the rates are quite low (and even lower for multiple suspensions) that plays a large role in the high proportion blacks and other disadvantaged groups make up of students suspended one or more times. And actions recommended in the Policy Statement to generally reduce preschool suspensions will tend to increase those proportions.

Another recent example where the government has focused on the high proportion a group makes up of persons experiencing some adverse outcome in analyzing discrimination issues may be found in the DOJ’s contentions regarding the disparate impact of police and court practices in Ferguson Missouri. Both the investigative report the DOJ issued in March 2015 and the suit it brought in February 2016 maintain that unduly aggressive policing practices and unjustifiably harsh court procedures cause blacks to make up much higher proportions of persons experiencing adverse interactions with police and courts than they make up of the city’s population. As explained above, modifying practices in ways that generally reduce those

adverse interactions will tend to increase, rather than reduce, the proportion blacks make up of persons experiencing them.

The report and the complaint also reflect the view that the continuation of existing practices, in light of the evident magnitude of their disproportionate effect, supports an inference that the practices are motivated by racial animus. Thus, the DOJ's mistaken perception about the relationship of the frequency of an outcome to measures of racial disparity creates not only a situation whereby following government guidance for reducing racial impact of a practice will tend to cause racial disparities generally, and the disparate impact of particular practices, to be deemed larger than otherwise would be the case. It also creates a situation whereby following such guidance will tend to strengthen evidence that practices are maintained for a discriminatory purpose.

A consent decree entered in the Ferguson case in April 2016 imposes on the city a variety of obligations regarding the monitoring of racial differences in the city's criminal justice system. Some of the subjects to be monitored are cast in terms of favorable outcomes and some are cast in terms of adverse outcomes. The decree also requires that when the city identifies a practice with a disparate impact, it shall attempt to identify alternative practices "with less of a disparate impact." In fulfilling these obligations, the city is unlikely to recognize that the generally more lenient police and court practices that are a theme of the decree will tend to result in reduced racial disparities for things cast in terms of favorable outcomes, but increased racial disparities for things cast in terms of adverse outcomes. That holds as well for the entities to be retained, at considerable expense to the taxpayers of Ferguson, to monitor the city's compliance with the decree.

Until the understanding of these issues changes dramatically, we can also expect that communities seeking to avoid actions of the type the DOJ has taken against Ferguson may unwittingly modify practices in ways that will have an effect that is the opposite of the intended effect. That holds both as to the percentage differences between rates at which minorities and whites experience adverse interactions within the criminal justice system (or the proportion minorities make up of persons experiencing those interactions) and as to the likelihood of avoiding action by the DOJ.

Employment Settings

Returning to the employment settings to which Uniform Guidelines specifically pertain, one might think that, whether or not the measurement of disparate impact in employment makes sense, at least there does not exist the anomaly whereby an entity that relaxes standards increase chances it will be sued for discrimination. But clarifying Questions and Answers accompanying the Guidelines provide that in circumstances where the favorable outcome is so common that it is impossible to violate the four-fifths rule, disparate impact may be measured in terms of percentage differences in the adverse outcome.

Thus, even with respect to an employment test there exists the possibility that once the disadvantaged group's pass rate reaches 80%, further lowering of cutoffs could be read as increasing the disparate impact of the test, as measured in terms of percentage differences in

failure rates. Irrespective of this feature of the Guidelines, however, disparities in testing outcome have often been cast in terms of percentage differences in failure rates. See, e.g., *Washington v. Davis*, 426 U.S. 229 (1976), where the Court addressed the disparate impact of the test at issue in terms of the fact that the black failure rate was four times the white rate. Then, as now, virtually no one understood that lowering the cutoff would tend to increase that difference.

But there also exists a range of non-testing situations where employment disparities are measured in terms of percentage differences in adverse outcomes. These include situations involving policies of refusal to hire persons with arrest or conviction records and policies of terminating employees for failure to meet performance or discipline standards. So far, however, those addressing such issues have failed to recognize that more narrowly tailored policies of refusal to hire for arrest or conviction records, or more lenient performance and discipline policies, will tend to increase percentage differences between the rates at which whites and minorities are adversely affected by those policies.

Appraising the Comparative Size of Disparate Impacts

Let us return to the question of whether relaxing a standard, and thereby reducing percentage differences in rates of meeting it and increasing percentage differences in rates of failing to meet it, reduces or increases the standard's disparate impact. In a situation where meeting or failing to meet the standard entirely dictates the ultimate outcome, there seems no rational basis for maintaining that the size of the disparate impact varies depending on the stringency of the standard. For example, where all who pass a test experience the desired outcome, and all those who fail the test do not, as would typically be the case with bar exams and other certification procedures as well as teacher competency and high school exit exams, the disparate impact of the test would seem unaffected by the cutoff. That seems also to hold with respect to things like standards for termination from a job for inadequate performance or misconduct and criteria for mortgage foreclosure.

Where the challenged criterion merely determines who will be subject to a further selection process, the situation becomes rather more complicated, and determinations will be affected by how persons from different groups fare in the further selection process. And there may be circumstances where a lower standard might be deemed to have less of a disparate impact (as in the case of a test cutoff to advance further in the employment process) and circumstances where a higher standard might be deemed to have less of a disparate impact (as in the case of the performance rating level that exempts employees from being subject to a reduction in force based on seniority). There will also be situations where it is very difficult to determine the contribution of meeting or failing to meet a criterion to ultimate outcomes or the effect of the stringency of the standard on differences in those outcomes. I treat these complicated issues, sketchily and with some uncertainty, in Section E of a 2013 University of Kansas School of Law faculty workshop paper titled "[The Mismeasure of Discrimination](#)." I also note, however, that the matter may be further complicated by Supreme Court precedent, in *Connecticut v. Teal*, 457 U.S. 440 (1982), holding that the disparate impact is determined at the point where the challenged standard limits the pool eligible for further consideration.

But such issues can only be examined with anything approaching rationality in settings where all participants understand that the two percentage differences tend to change in opposite directions as a standard is raised or lowered, as well as the ways other measures tend to be affected by the frequency of an outcome, and, equally important, that a measure that changes solely because the frequency of an outcome changes cannot effectively quantify the differences in the circumstances of two groups reflected by their favorable or adverse outcome rates. Participants must also understand that, while one may effectively (if imperfectly) quantify such differences when one knows the actual outcome rates for the groups being compared, it is impossible to do so based solely on information on the proportion a group makes up of persons potentially experiencing an outcome and the proportion it makes up of persons actually experiencing the outcome. See Section C of the Kansas Law paper. We may be many decades away from a general recognition of such things, even in the scientific community. But in a litigation it is at least theoretically possible that all parties, and the court, can be made to understand such things as they bear on the particular matter at issue.

It is difficult even to guess whether many courts understanding the relevant issues would regard the disparate impact doctrine, or its less discriminatory alternative element, to be unconstitutionally vague. It is difficult even to estimate the role the fact that these matters have heretofore been unknown to the Congress enacting the subject statutes, or to the agencies and courts enforcing them, would play in such determination. It is similarly difficult to assess whether courts fully understanding the matter will be better disposed toward a vagueness argument than courts that only dimly understand the matter, with or without recognizing the limits of their understanding.

But, whether or not there will exist a basis to void disparate impact components of statutes for vagueness once the issues are fully understood, there is a compelling need for those issues to be better understood than they are now.

For a fuller explanation of the patterns by which standard measures of demographic differences tend to be affected by the frequency of an outcome, and the implications of the failure to understand the patterns by those examining group differences in the law and the social and medical sciences, see, in addition to the ASA letter and Kansas Law paper already mentioned, my "[Race and Mortality Revisited](#)," *Society* (July/Aug. 2014), "[The Perverse Enforcement of Fair Lending Laws](#)," *Mortgage Banking* (May 2014), [amicus curiae](#) brief in *Texas Department of Housing and Community Development, et al. v. The Inclusive Communities Project, Inc.*, Supreme Court No. 13-1731 (Nov. 2014), and the methods workshops collected [here](#). For a discussion of the entirely unsatisfactory manner in which the National Center for Health Statistics has dealt with its recognition that percentage differences in favorable health and healthcare outcomes and percentage differences in the corresponding adverse outcomes tend to change in opposite directions as health and healthcare generally improve, see my forthcoming "The Mismeasure of Health Disparities," *Journal of Public Health Management & Practice* (July/Aug. 2016).