Portfolio Media. Inc.
860 Broadway, 6th Floor, New York, NY 10003
Phone: +1 646 783 7100 | Fax: +1 646 783 7161
www.law360.com | customerservice@law360.com

How Bad Data Analyses Can Sabotage Discrimination Cases

Law360, New York (December 15, 2016)
When Mark Twain popularized the phrase, "there are three kinds of lies: lies, damned lies, and statistics," he was throwing out the baby with the bathwater.


Mike Nguyen

When utilized properly, statistics can be a highly effective tool in supporting a legal argument. Conversely, statistics that are misunderstood or misinterpreted can quickly sink a case. In discrimination cases, as in so many other matters, it is critical to understand both the benefits and limitations of using data to support an argument.

We're going to ask our friends Serena Williams and Derek Jeter to help us make the case. But first, some background, and then some recommendations for attorneys involved in these types of cases.

Over the years, statistical data and analyses have been used in courtrooms as evidence of race-, gender- or age-based discrimination. As documentary evidence of discriminatory practices is rarely available, courts often rely on additional evidence, including statistical analyses, to evaluate whether the process or practice in question may have been discriminatory. However, the courts have been cautious in using statistical evidence to evaluate whether discrimination has in fact occurred, realizing that the statistical evidence itself may contain bias or yield misleading results.

This article discusses a common flaw with statistical analyses that can have a material effect on the outcomes.

Inclusive Communities Opens the Door

In the Texas Department of Housing and Community Affairs v. Inclusive Communities Project Inc. Fair Housing Act case (Inclusive Communities), the importance of sound statistical analysis in race-, gender- or age-based discrimination cases was amplified. In its 2015 decision, the U.S. Supreme Court addressed whether the Fair Housing Act recognizes disparate impact claims,[1] in which a plaintiff can establish liability, without proof of intentional discrimination, if a particular business practice has a disproportionate effect on specific groups of individuals and if the practice is not based on sound business considerations.

The court found that, based on its language and purpose, the Fair Housing Act provides for disparate impact claims. However, the court also imposed important limitations "to protect potential defendants against abusive disparate-impact claims."[2] Here's where statistical analysis comes in.

Limitations of Statistical Analyses and "Simpson's Paradox"

It is important to recognize that statistical analyses, including regression analyses, cannot perfectly determine whether differences in the outcomes for certain groups of individuals are due to discrimination. For example, the large body of academic literature on race-based discrimination in mortgage lending shows that one cannot look at outcomes in a regression model and infer race-based discrimination without having controlled for all factors that explain differences in mortgage credit decisions and without reviewing individual loan files.[3] There are many variables that affect mortgage interest rates paid by borrowers that typically are not available for the purposes of a regression analysis. These include factors that may be nonquantifiable, such as a borrower's experience.

When using statistical analyses to test for discrimination in mortgage lending and other practices, experts often aggregate data in order to draw conclusions about particular groups as a whole. However, experts need to be cautious not to offer misleading conclusions that can result from such aggregation. A classic pitfall involves something known as "Simpson's Paradox."

Simpson's Paradox is a statistical concept that refers to the reversal of statistical relationships when two or more groups of data are combined. In a Simpson's Paradox situation, an analysis of aggregated data can produce a statistically significant result[4] even though the individual elements included in the aggregation may not be statistically significant.

A famous example of Simpson's Paradox involves claims of gender bias in graduate admissions at the University of California at Berkeley. Researchers analyzing aggregate graduate school admissions data across all graduate departments at the university found that male applicants were more likely to be admitted than female applicants, and that the difference was so large that it was unlikely due to chance.

However, each of Berkeley's departments maintained their own policies as to graduate admissions and made decisions independent of the other departments. When the analysis was conducted for each individual decision maker (i.e., the individual department), only four departments had a statistically significant bias in favor of males, while six departments actually had a statistically significant bias in favor of females — even though the aggregate analysis had suggested a statistically significant bias in favor of males across the entire university. The underlying cause of this aggregate result was that more women applied to the most competitive departments with the highest demand and the lowest admissions rates. In other words, women rejected from the highly competitive departments were overrepresented in the aggregate population.

To understand how the math in a Simpson's Paradox situation works, consider a simple example from Serena Williams. Imagine she is playing in the final of the Law360 Open against her sister Venus. Venus wins a total of 14 games and Serena only 12. Venus wins, right?

Not necessarily, because of the structure of the game. In women's tennis, the winner must take two of three sets. So while Venus won more games, did she win two sets? In our example, Venus wins the first set 6-0. But she loses the next two sets, each by the score of 6-4. So while Serena won fewer games (12 to her sister's 14), she wins the match (two sets to her sister's one). (See Table 1.)

Table 1

	Set 1	Set 2	Set 3	Games Won
S. Williams	0	6	6	12
V. Williams	6	4	4	14

Another example of Simpson's Paradox can be seen in comparing the batting averages of players in professional baseball. Consider this question:

Is it possible for one player to hit for a higher batting average than another player during a given year, and to do so again during the next year, but to have a lower batting average when the two years are combined? Surprisingly, the answer is "yes."

This can occur when there are large differences in the number of at-bats between the years. Ken Ross, a professor emeritus at the University of Oregon and baseball enthusiast, notes in a 2004 book on baseball statistics that in both 1995 and 1996, Derek Jeter of the New York Yankees had a lower batting average for each season than David Justice, then of the Atlanta Braves. (See Table 2.)

Table 2

	1995		1996	
	Hits/At-bats	Avg.	Hits/At-bats	Avg.
Derek Jeter	12/48	0.250	183/582	0.314
David Justice	104/411	**0.253**	45/140	**0.321**

Combining the two years, however, Derek Jeter had a higher batting average overall. The paradox here results from the fact that the majority of Derek Jeter's at-bats came in 1996, with a higher batting average (0.314), while the majority of David Justice's at-bats came in 1995, with a lower batting average (0.253). The combined figures push the two-year average in Derek Jeter's favor. (See Table 3.)

Table 3

	1995		1996		Combined	
	Hits/At-bats	Avg.	Hits/At-bats	Avg.	Hits/At-bats	Avg.
Derek Jeter	12/48	0.250	183/582	0.314	195/630	**0.310**
David Justice	104/411	**0.253**	45/140	**0.321**	149/551	0.270

Application to a Discrimination Case

If not examined carefully, statistics in the courtroom can be equally misleading. In a case in which Analysis Group provided expert analysis, a mortgage lender was accused of implementing a "discretionary pricing policy" under which mortgage brokers were given discretion in pricing mortgage loans. Plaintiffs' expert, relying on regression analysis of aggregated data, alleged that this policy caused African-American borrowers to receive higher annual percentage rates (APRs) for mortgage loans than similarly situated white borrowers.

However, if in fact there was a discretionary pricing policy, and that policy had a disparate impact on African-American borrowers, then the disparate impact caused by the policy should be observed consistently across the various mortgage brokers who applied it. If such disparate impact was not observed consistently across mortgage brokers, that would suggest that loan pricing was the result of individualized decision-making by the brokers rather than the result of a common policy applied across the entire borrower population.

When Analysis Group's expert (working on behalf of the defendant) disaggregated the mortgage data and analyzed it at the individual broker level, it became clear that the statistical evidence did not reflect a commonly applied "discretionary pricing policy" that resulted in higher loan prices for African-American borrowers as a group. In fact, when the top mortgage brokers were studied individually, the plaintiffs' expert's own regression model revealed that, for certain brokers, African-American borrowers received APRs that were lower than the APRs received by similarly-situated white borrowers.

Key Takeaways: Understand the Strengths and Limitations of Data in Discrimination Cases

The Supreme Court's decision in Inclusive Communities suggests that statistical analyses may play a larger role in discrimination cases because disparate impact alone may be sufficient to demonstrate that discrimination has occurred. For attorneys involved in these types of cases, there are a few important takeaways:

Statistical analysis can be a highly effective tool in supporting a legal argument.
It is critical to understand the assumptions upon which statistical tests and conclusions are based.
Statistical tests that are not done properly can produce misleading results.
And finally, if you're having a good year, try to get more at-bats.

Mike Nguyen is a Vice President in the Los Angeles office of Analysis Group, and a former wide receiver for the UCLA Bruins (1991 – 1995).

Endnotes

[1] The theory of disparate impact holds that practices in employment, housing, or other areas may be considered discriminatory and illegal if they have a disproportionate "adverse impact" on persons in a protected class. Disparate impact is different from disparate treatment, which involves discriminatory intent or motive. Both disparate impact and disparate treatment are considered forms of discrimination.

[2] For example, it emphasized that the plaintiff has the burden "to establish a 'robust' causal connection between the challenged practice and the alleged disparities." http://www.law360.com/articles/794007/us-supreme-court-impact-on-mass-disparate-impact-claims.

[3]]See, e.g., Paul Calem and Stanley Longhofer, Anatomy of a Fair Lending Exam: The Uses and Limitations of Statistics 24 J. REAL ESTATE FIN. 207 (2002).

[4] A statistically significant result is a result that is unlikely to have occurred purely by chance.