Statistics, civil rights and the US Supreme Court: A cautionary tale

A sidelining of statistical evidence in civil rights cases has been 30 years in the making.
By **Daniel Kiefer**, **Jerome Sacks** and **Donald Ylvisaker**

The nine justices of the Supreme Court of the United States have the last word on appeals of rulings in the federal court system, and on those from individual state courts that involve federal law. In 2011, the Court rejected a class-action discrimination lawsuit alleging large gender differences in pay and promotions at Wal-Mart Stores, America's largest private business. The decision in *Wal-Mart* v. *Dukes* was based on the Court's strained but arguably defensible construction of technical requirements for certification of a federal class action; with this the Court discarded statistical evidence suggestive of widespread sex discrimination at Wal-Mart. The consequences: scant legal recourse for its women employees, and a substantial barrier to class-action discrimination suits in general.

Two years later in *Shelby County* v. *Holder*, the Court eliminated a key protection against voter discrimination

Image: bbourdages/iStock/Thinkstock

A note on disparate impact

A disparate impact claim inevitably has statistics at its core, and that is our focus. However, a disparate treatment theory ("I was treated unfairly because of my race") is also available to support a claim of discrimination. Disparate treatment claims also use statistical evidence, but it is not sufficient: a claimant using this theory must show intent to discriminate.

that, historically, had featured barriers to the voting of black Americans in the southern states. The Court held that the protection of the vote then in place was based on outmoded data; in so doing, it ignored voluminous data assembled by Congress that showed continuing need for voter protection. In the states previously under voting change restrictions, the decision prompted a spate of new laws limiting access of voters to the polls, laws that especially affected minorities.

The Court's treatment of statistical evidence in discrimination cases has grown increasingly hostile, dating back to the Reagan era; *Wal-Mart* and *Shelby* represent merely the latest wave of a tsunami 30 years in the making. How and why did we end up here? The answer lies in tracing a succession of critical Supreme Court decisions that, by now, speak to a judicial view of statistical evidence born of something other than healthy analytical scepticism.

Daniel Kiefer is a writer and attorney in Los Angeles

Jerome Sacks is director emeritus of the National Institute of Statistical Sciences

Donald Ylvisaker is professor emeritus at the University of California at Los Angeles

Griggs and the birth of disparate impact

In the mid-1960s, the civil rights movement led the American federal government to respond to the country's seemingly intractable discrimination problems with a series of landmark laws. At the time, overt discrimination ("blacks need not apply") was commonplace. The Civil Rights Act (CRA) of 1964 prohibited discrimination in employment, schools and public facilities based on race, colour, religion, sex, or national origin. Inadequacies of the CRA led to more specific legislation, including the Voting Rights Act (VRA) in 1965 and the Fair Housing Act (1968). The new laws, in effect, forced discrimination to take cover; subtler methods of bias – covert means to a comparable end ("high school diploma required") – replaced blatant racism.

A new challenge to the civil rights movement thus came into focus: a more nuanced bias seen, for example, in business practices designed to hide discrimination. When a suit alleging covert discrimination, *Griggs* v. *Duke Power*, reached the Court, it held unanimously that Title VII of the CRA "proscribes not only overt discrimination, but also practices that are fair in form, but discriminatory in operation".

Griggs thereby established the doctrine of disparate impact (see "A note on disparate impact"). After *Griggs*, protected workers (in this case blacks) need not show intent to discriminate, only that an employer's practice had a disparate impact on the class. If a claimant did so, the burden shifted to the employer to show the business necessity of the practice. The key use of statistical evidence in showing disparity was only implicit in *Griggs*; it became explicit in later decisions in the 1970s.

Griggs had sweeping consequences. In its wake, the disparate impact principle was acknowledged by the lower courts in rulings on age and housing discrimination cases. It also came to be applied in voting rights cases when Congress, in 1982, amended the VRA so that its Section 2 specifically barred electoral practices that resulted, intentionally or not, in denying a racial or language minority an equal opportunity to participate in the political process. In *Thornburg* v. *Gingles* (1986), the Court not only upheld the amended VRA, but also endorsed the use of statistical evidence to establish racially polarised voting and consequent inability of a cohesive minority to elect candidates of their choice.

The enforcement of the keystone civil rights laws of the 1960s therefore hinged heavily on statistical evidence. Practices and policies with discriminatory effects, intentional or not, were barred, and statistics were used to establish these effects in the absence of direct evidence of intent. But as the political climate changed over the 25 years following enactment of these laws, so did the composition of the Court. In 1986, Chief Justice Warren Burger retired. The elevation of William Rehnquist to Chief Justice and the appointment of Antonin Scalia to the Court highlighted a period of increased judicial conservatism and polarisation at the Court that weakened civil rights laws, and – not coincidentally, and of particular concern here – brought a marked devaluation of statistical evidence.

McCleskey: The tide turns

The first blow struck by the new Court came in a suit charging racial bias in death penalty sentencing. In *McCleskey* v. *Kemp* (1987), a challenge to the State of Georgia's death penalty law, McCleskey, a black man convicted of killing a white victim, claimed that the capital sentencing process was unconstitutional on the basis of the Eighth ("cruel and unusual punishments") and Fourteenth ("equal protection of the laws") Amendments. He relied upon an extensive statistical investigation of 2482 death penalty sentencing cases in the state of Georgia between 1973 and 1978.[1] It found, as one example, that after consideration of 230 mitigating circumstances and accounting for the 39 most significant of them through multivariate regression, the odds of receiving a death sentence were 4.3 times greater for those whose victims were white than for those whose victims were black.

Despite this and other significant disparities in the imposition of the death sentence based on both the victims' and the defendant's race, the Court ruled, by a 5–4 vote, that evidence of discrimination against a group is not a violation of an individual member's constitutional rights: "The statistics do not prove that race enters into any capital sentencing decisions or that race was a factor in petitioner's case."

With this ruling, the Court set intentional discrimination as the constitutional standard for death penalty appeals, setting aside statistical evidence of racial disparity in sentencing. The virtual impossibility of showing intentional discrimination by prosecutors and/or juries has spread the effect of *McCleskey* beyond capital punishment, stifling suits charging discrimination in other criminal justice settings. For example, studies have shown that blacks receive longer

sentences in drug cases than whites, but *McCleskey* places an insurmountable bar to rectifying that in the courts.

A year later, in *Watson* v. *Fort Worth Bank and Trust*, the Court extended the reach of employment discrimination law to subjective employment practices. As it broadened the reach of disparate impact, the ruling carried some extra baggage. The Court held that the plaintiff must do more than "show that there are statistical disparities in the employer's work force. The plaintiff must begin by identifying the specific employment practice that is challenged." It went on to say: "Once the employment practice at issue has been identified, causation must be proved; that is, the plaintiff must offer statistical evidence of a kind and degree sufficient to show that the practice in question has caused the exclusion of applicants for jobs or promotions because of their membership in a protected group."

In 1989, in a 5–4 decision, *Wards Cove* v. *Atonio* reaffirmed *Watson* as law, and effectively revoked the *Griggs* disparate impact principle. For one, the Court held that "the ultimate burden of proving that discrimination against a protected group has been caused by a specific employment practice remains with the plaintiff at all times". The opinion also clarified the new "causation" requirement, holding that plaintiffs must "demonstrate that the disparity they complain of is the result of one or more of the employment practices that they are attacking here, specifically showing that each challenged practice has a significantly disparate impact on employment opportunities for whites and nonwhites". This "practice-by-practice" requirement, and its implicit call for dissection of the evidence data, led Justice Harry Blackmun to charge the majority with taking a major step backwards in the battle against racial discrimination by requiring "practice-by-practice statistical proof of causation, even where, as here, such proof would be impossible".

In 1991 Congress reacted to the Court's perceived assault on civil rights legislation by amending the CRA to incorporate the language of *Griggs*, reaffirming disparate impact and restoring the employer's business necessity burden. Yet here "[t]he mere existence of a statistical imbalance in an employer's workforce on account of race, color, religion, sex, or national origin is not alone sufficient to establish a prima facie case of disparate impact violation". The amended Title VII retained the "practice-by-practice" requirement from *Wards Cove* without explicit guidance on how plaintiffs would prove causation. In other efforts to provide protection to victims of discrimination, the Act provided for compensatory damages and the right to trial by jury.

Wal-Mart and beyond

The Act's intended effects received mixed reviews from the courts. While there was an uptick in charges filed with the Equal Employment Opportunity Commission and several large class-action settlements were obtained, the possibility of monetary damages and/or a jury trial complicated the issue of class certification. Several district courts had addressed this added complexity when, in 1998, the first court of appeals ruling on it found that these changes "are not inconsequential

Wal-Mart data: an example

The study of Wal-Mart promotion and pay data in Gastwirth *et al.*,[2] for example, finds promotion data for the regions in *Dukes* v. *Wal-Mart* consistent with a system in which the odds of a female employee receiving a promotion to one of the four managerial positions in question are about 70–80% of those of a male, as well as statistically strong evidence of a pattern of underpayment of female employees relative to comparable males.

... [i]n the class action context". In 2001, when 1.5 million female employees at Wal-Mart filed a class-action sex discrimination lawsuit alleging inequities in pay and promotion, certification of the class became the central issue. The district court certified the class and the appellate courts affirmed the certification. In 2009, Wal-Mart appealed these rulings to the Supreme Court.

In *Wal-Mart* v. *Dukes* (2011) the Court ruled that the case[s] should not have been certified as a class action because the members of the class did not have a claim in common ("commonality" is a requirement of class certification). In the 5–4 decision Justice Scalia wrote: "Wal-Mart's 'policy' of allowing discretion by local supervisors over employment matters ... is just the opposite of a uniform employment practice that would provide the commonality needed for a class action; it is a policy against having uniform employment practices." And, citing the practice-by-practice requirement from *Watson* and *Wards Cove*, Scalia identified the specific practices at Wal-Mart as the actions of individual managers. Since individual managers are certain to act differently, commonality is absent.

A succession of critical Supreme Court decisions speak to a judicial view of statistical evidence born of something other than healthy analytical scepticism

Statistical analyses indicating gender discrimination at Wal-Mart (see "Wal-Mart data: an example") thus became irrelevant because the case failed at the class certification stage. Through the illogic of "discretion is a non-policy policy", the Court could turn a blind eye to the work environment of the women employees of Wal-Mart. The ruling erected a high barrier for any class-action discrimination suit under Title VII of the CRA.

That statistical evidence posed no barrier to forming the Court's opinion was on full display in *Shelby County* v. *Holder* (2013). Here the majority addressed a key provision, Section 5, of the VRA that required specified jurisdictions (mostly in the

southern states) to have proposed changes in voting practices or procedures "precleared". Since 1965, this law has had a prominent role in ensuring increased minority access at the polls and in electing minority candidates. When the VRA was renewed in 2006, Congress undertook an extensive study of preclearance history. The more than 15 000 pages of this record showed the many advances made by minorities, especially African-Americans, as well as voluminous evidence of efforts to institute practices detrimental to the voting rights of minorities and of the denial of clearance by the Justice Department.

Relying on these detailed factual findings, Congress voted overwhelmingly to extend the preclearance process, and this extension was signed into law by President George W. Bush. Despite this, the Court ruled, by a 5–4 majority, that the criteria subjecting jurisdictions to preclearance in Section 4 of the VRA were unconstitutional because they were based on outmoded data, effectively nullifying Section 5.

Chief Justice John Roberts, writing for the majority, found that the data Congress gathered indeed showed that the VRA had been successful in increasing minority participation. Ignoring the evidence of ongoing discrimination, including the large number of recent objections to proposed changes in voting practices in the preclearance process, Roberts concluded that the success of Section 5 of the VRA and "current conditions" implied it was no longer needed.

As Justice Ruth Bader Ginsburg observed in dissent: "[T]he Court strikes [the] coverage provision because, in its view, the provision is not based on 'current conditions' ... It discounts, however, that one such condition was the preclearance remedy in place in the covered jurisdictions, a remedy Congress designed both to catch discrimination before it causes harm, and to guard against return to old ways."

Only a few hours after the *Shelby County* decision, Texas implemented a restrictive voter ID law. Other formerly covered states followed suit with similar restrictive voting measures. Whether those new restrictions will affect the 2016 elections is an open question.

What can we draw from this story? One view is that statistics has merely been collateral damage in the political/ideological wars over civil rights. Supreme Court decisions have long been recognised as driven in part by politics, but the sidelining of statistical evidence in *McCleskey* and *Wal-Mart*, its downgrading in *Watson*, and its selective usage in *Shelby County* introduce a more troublesome element. A jurisprudence that is cavalier or mistaken about statistical evidence hardly serves a time when data and their analyses and interpretations are so embedded in our lives. ■

References

1. Baldus, D. C., Pulaski, C. and Woodworth, G. (1983) Comparative review of death sentences: An empirical study of the Georgia experience. *Journal of Criminal Law and Criminology*, **74**(3), 661–753.

2. Gastwirth, J. L., Bura, E. and Miao, W. (2011) Some important statistical issues courts should consider in their assessments of statistical analyses submitted in class certification motions: Implications for *Dukes* v. *Wal-Mart*. *Law, Probability and Risk*, **10**, 225–263.

In the Court's own words

***Teamsters* v. *US* (1977)**
"We have repeatedly approved the use of statistical proof ... in jury selection cases ... Statistics are equally competent in proving employment discrimination." (Justice Potter Stewart for the majority)

***Hazelwood* v. *US* (1977)**
"Where gross statistical disparities can be shown, they alone may in a proper case constitute prima facie proof of a pattern or practice of discrimination." (Justice Potter Stewart for the majority)

***Thornburg* v. *Gingles* (1986)**
" ... we would hold that the legal concept of racially polarized voting ... refers only to the existence of a correlation between the race of voters and the selection of certain candidates. Plaintiffs need not prove causation or intent in order to prove a prima facie case of racial bloc voting." (Justice William Brennan for the Court)

***McCleskey* v. *Kemp* (1987)**
"The likelihood of racial prejudice allegedly shown by the [Baldus] study does not constitute the constitutional measure of an unacceptable risk of racial prejudice." (Justice Lewis Powell for the majority)

"Close analysis of the Baldus study, however, in light of both statistical principles and human experience, reveals that the risk that race influenced McCleskey's sentence is intolerable by any imaginable standard." (Justice William Brennan, dissenting)

"I do not share the view, implicit in the opinion, that an effect of racial factors upon sentencing, if it could only be shown by sufficiently strong statistical evidence, would require reversal." (Justice Antonin Scalia, joining the majority, in a Memorandum to Conference)

***Watson* v. *Fort Worth Bank and Trust* (1988)**
"Our formulations ... have consistently stressed that statistical disparities must be sufficiently substantial that they raise such an inference of causation ... Nor are courts or defendants obliged to assume that plaintiffs' statistical evidence is reliable." (Justice Sandra Day O'Connor for the Court)

***Wal-Mart Stores* v. *Dukes* (2011)**
"Even if it [statistics] established (as it does not) a pay or promotion pattern that differs from the nationwide figures or the regional figures in all of Wal-Mart's 3400 stores, that would still not demonstrate that commonality of issue exists. Some managers will claim that the availability of women, or qualified women, or interested women, in their stores' area does not mirror the national or regional statistics. And almost all of them will claim to have been applying some sex-neutral, performance-based criteria – whose nature and effects will differ from store to store." (Justice Antonin Scalia for the majority)

***Shelby County* v. *Holder* (2013)**
"Congress compiled thousands of pages of evidence before reauthorizing the Voting Rights Act. ... Regardless of how to look at the record, however, no one can fairly say that it shows anything approaching the 'pervasive,' 'flagrant,' 'widespread,' and 'rampant' discrimination that faced Congress in 1965." (Chief Justice John Roberts for the majority)

"The Court makes no genuine attempt to engage with the massive legislative record that Congress assembled. ... Without even identifying a standard of review, the Court dismissively brushes off arguments based on 'data from the record,' and declines to enter the 'debat[e about] what [the] record shows.' ... One would expect more from an opinion striking at the heart of the Nation's signal piece of civil-rights legislation." (Justice Ruth Bader Ginsburg, dissenting)