Age discrimination claims are on the rise, both in volume of cases filed and size of verdicts. Given the current economic landscape, widespread layoff announcements, and the aging boomer generation, that trend seems unlikely to abate anytime soon. Meanwhile, the Supreme Court last term granted certiorari in *Adams v. Florida Power Corp.*, to resolve a split among the circuits over the viability of the disparate impact theory of liability (as opposed to the disparate treatment theory) in cases brought under the federal Age Discrimination in Employment Act of 1967 (“ADEA”). Commentators had predicted that the Supreme Court would likely hold that disparate impact was not a viable theory for age discrimination claims. Because disparate impact claims are almost always based on statistical evidence, such a ruling could also call into question the relevance of statistical analyses of the impact of employment decisions on employees in age discrimination cases.

The viability of disparate impact claims under the ADEA remains an open question because the Supreme Court dismissed the appeal in *Adams* without deciding it. However, even if the Supreme Court ultimately holds that ADEA claims cannot be based on the disparate impact theory of liability, statistical evidence will continue to play an important role in age discrimination litigation. Employers would be wise to continue to scrutinize their employment policies and decisions, especially when planning large-scale reorganizations or reductions in force (RIFs), to ensure they do not disproportionately disadvantage older workers. This article discusses the viability of the disparate impact theory in cases alleging age discrimination, the use of statistical evidence in disparate treatment cases, and the practical cumulative effect of multiple claims of discrimination against the same employer in a single trial.
I. The Disparate Impact Theory under the ADEA

In a disparate treatment case, an employee must show that his employer treated him less favorably on account of his membership in a protected class like race, gender, or age. The employee is required to prove through direct or circumstantial evidence that the discrimination was intentional. In contrast, the disparate impact theory does not require proof of discriminatory intent. Rather, a facially neutral employment policy or practice that has a disproportionately adverse impact on the protected class of employees constitutes unlawful discrimination unless the employer establishes business necessity. Because proof of intent is not required, employees typically attempt to prove disparate impact cases through the use of statistical evidence.

The disparate impact theory has long been recognized as a valid method of proving race or gender discrimination under title VII of the Civil Rights Act of 1964. Federal courts have been sharply divided, however, on whether the ADEA imposes disparate impact liability without proof of an intent to discriminate. The Supreme Court expressly left that question open in Hazen Paper Co. v. Biggins although the Court’s analysis of the disparate treatment claim has led other courts and commentators to doubt the viability of disparate impact age discrimination claims.

The plaintiffs in Adams v. Florida Power were laid off during a number of reorganizations between 1992 and 1996, following the deregulation of the utility industry. They claimed that the reorganizations had a disparate impact on older workers, but the district court and a court of appeals held that the disparate impact theory was unavailable under the ADEA. The Eleventh Circuit relied on section 623(f)(1) of the ADEA, which permits employers to take action “based on reasonable factors other than age.” The court noted that
title VII, which does impose disparate impact liability, contains no provision similar to section 623(f)(1), but that the Equal Pay Act, which does not permit disparate impact claims, contained a provision analogous to section 623(f)(1). The court also relied on the legislative history of the ADEA and on the rationale of the Supreme Court’s decision in Hazen Paper in rejecting the disparate impact theory. In a concurring opinion, Judge Barkett disagreed with the majority’s view of disparate impact liability under the ADEA, but she concurred in the outcome because she believed that the plaintiffs had not sufficiently alleged a disparate impact claim.

In Adams v. Florida Power, the Eleventh Circuit joined the U.S. Courts of Appeals for the First, Seventh, and Tenth Circuits in holding that the ADEA does not impose liability for disparate impact claims. These circuits had cited the same reasons as Adams. Two of them, the First and Tenth Circuits, also noted that Congress had amended both title VII and the ADEA in 1991 but codified the disparate impact theory only for title VII. The courts reasoned that Congress’ conspicuous failure to mention the disparate impact theory in ADEA amendments signaled a legislative intent to exclude disparate impact liability under the ADEA. The First Circuit acknowledged the pitfalls of gleaning legislative intent from the conduct of a subsequent Congress, but the court found the 1991 amendments to be “a useful source of guidance.” Even if the 1991 amendments do not establish congressional intent in enacting the ADEA in 1967, however, the amendments at least dispel the notion that subsequent congresses, by their inaction, tacitly approved those pre-Hazen decisions that authorized disparate impact liability under the ADEA.

Outcomes in other circuits have been mixed. The U.S. Courts of Appeals for the Second, Eighth, and Ninth Circuits announced that they will permit ADEA suits based on the disparate impact theory. For the most part, these circuits have not engaged in any detailed analysis of the
issue, resting their decisions primarily on pre-\textit{Hazen} precedent and the general textual similarities between the ADEA and title VII.\textsuperscript{24} Courts in the Second Circuit have rebuffed attempts by plaintiffs to expand disparate impact theory to encompass statistical differences between subgroups protected by the ADEA, such as employees age fifty and over.\textsuperscript{25} The Second Circuit has also rejected plaintiffs’ attempts “to circumvent the subjective intent requirement in any disparate treatment case” by challenging the same employment decision under both the disparate impact and treatment theories.\textsuperscript{26}

The U.S. Court of Appeals for the D.C. Circuit has “assume[d] without deciding that disparate impact analysis applies to age discrimination claims.”\textsuperscript{27} The U.S. Courts of Appeals in the Third and Sixth Circuits have expressed serious doubts about the validity of disparate impact liability under the ADEA.\textsuperscript{28}

On December 3, 2001, the Supreme Court granted certiorari in \textit{Adams} to resolve this split of authority.\textsuperscript{29} Amicus curiae briefs were filed on behalf of no fewer than fifteen organizations.\textsuperscript{30} The Court heard oral arguments on March 20, 2002, but dismissed the appeal just twelve days later, stating that certiorari had been “improvidently granted.”\textsuperscript{31} Some of the questions the Court posed during oral argument suggest that the Court may have come to agree with Judge Barkett’s view that the employees had not adequately alleged a disparate impact claim and had failed to identify a facially neutral policy that allegedly had a disparate impact on older workers.\textsuperscript{32} Having sidestepped the issue twice, the Supreme Court likely will accept certiorari in a future case presenting the ADEA disparate impact issue. \textit{Hazen Paper} suggests that when the Court eventually reaches the issue, it will reject the disparate impact theory as a method of proving age discrimination under the ADEA.\textsuperscript{33}
Whatever the outcome, however, for three reasons employers should continue to monitor the statistical impact of their employment decisions and policies upon employees age forty and older. First, the disparate impact theory remains viable under a number of state statutes prohibiting age discrimination. Second, courts routinely permit the introduction of statistical evidence in support of disparate treatment cases. Third, the mere fact that multiple claims of age discrimination are tried together in a single case, regardless of the individual merit of any one of them, can sway a jury in the plaintiffs’ favor and lead to significant jury awards against employers. In such cases, employers may counter this prejudicial effect by introducing their own statistical evidence.

II. Disparate Impact under State Age Discrimination Statutes

By the time Congress enacted the ADEA, many state statutes had already prohibited employment discrimination based on age. Today such laws exist in virtually every state. The ADEA specifically provides for concurrent state and federal enforcement of age discrimination laws, and it requires claimants to initiate state administrative proceedings before filing suit under the ADEA. Many states permit age discrimination claims based on the disparate impact theory, including California, Maine, Massachusetts, Michigan, Minnesota, Nebraska, Tennessee, Vermont, Washington, and West Virginia. In fact, California and Minnesota recently amended their age discrimination laws expressly to authorize disparate impact claims. As a result, regardless of whether disparate impact liability under the ADEA is eliminated, plaintiffs around the country will be able to base age discrimination claims on the disparate impact theory.

III. Statistical Evidence in Disparate Treatment Cases

Even if the Supreme Court ultimately holds that disparate impact is not a viable theory of liability for claims brought under the ADEA, employees will continue to use statistical evidence
in support of ADEA disparate treatment claims. Plaintiffs are seldom able to adduce direct
evidence that their employers intended to discriminate against them on account of their age. For
this reason, plaintiffs are permitted to introduce circumstantial evidence of intent in disparate
treatment cases, and frequently such evidence takes the form of statistics.\textsuperscript{38} The Supreme Court
has specifically endorsed the use of statistical evidence in support of disparate treatment age bias
claims in both individual and class-wide pattern-and-practice discrimination cases.\textsuperscript{39} Some
courts have even held that a plaintiff can establish a prima facie case of disparate treatment on
the basis of statistical evidence alone.\textsuperscript{40} There is little practical difference, then, between the
disparate impact theory, in which a plaintiff uses workforce statistics to show that a facially
neutral policy disproportionately disadvantages older workers, and the disparate treatment
theory, in which a plaintiff uses statistics as circumstantial evidence of a systemic pattern or
practice of intentional age discrimination.

Statistical evidence is particularly troublesome for employers defending discrimination
claims, for several reasons. First, statistical evidence likely precludes summary judgment for the
employer.\textsuperscript{41} Employers therefore must suffer the disruption and expense of lengthy and
extensive discovery and a jury trial even though the case may have little merit simply because
the plaintiff’s expert is able to present the statistical data in a manner that suggests discrimination
in employment decision-making.

Second, once the case goes to trial, employers often find it difficult to counter the
plaintiff’s statistical evidence. Juries are ill-equipped to evaluate the relative validity of two
competing statistical methods, and judges are unlikely to assist them in any meaningful way.
Because juries generally sympathize with age discrimination plaintiffs,\textsuperscript{42} they are likely to favor
the plaintiff’s statistics. In addition, juries usually prefer a simple, easy-to-grasp factual theory
to a complex and intellectually rigorous one. This also tends to benefit plaintiffs in age bias cases. The plaintiffs’ statistician will often adopt a simplistic approach without considering significant variables other than age that might account for perceived discrepancies in treatment of older workers.

Even after the Supreme Court’s decisions in *Daubert* and *Kumho Tire*, which announced more stringent standards for the admission of expert testimony in federal courts, courts have declined to bar plaintiffs’ evidence that is not “rigorously statistical,” viewing such flaws in methodology as going to the weight and not the admissibility of the statistical evidence. As the Supreme Court has said, “Normally, failure to include variables will affect the analysis’ probativeness, not its admissibility.” Further, courts generally agree that the appropriate statistical methodology to use in discrimination cases must be determined on an *ad hoc*, case-by-case basis. In addition, courts have been more liberal in accepting statistical evidence in individual disparate treatment claims to show that an employer’s justification for adverse action is merely a pretext for discrimination even though those same statistics fail to show systemic disparate treatment.

The Supreme Court and the lower federal courts have endorsed the standard deviation test as a method of gauging whether apparent discrepancies are statistically significant. Under this test, a variance of more than two standard deviations from the expected outcome gives rise to an inference that the deviations were not the product of chance and that age may have played a role in the employer’s decision-making. For example, assume an employer has a total workforce of 200 employees, half of whom are age forty or older. The employer lays off fifty employees, forty of whom are age forty or older and only ten of whom are under age forty. All other things being equal, one would have expected half of the employees selected for layoff
(twenty-five) to be under age forty, and half to be age forty or older. Under the standard deviation test, discrimination might be inferred because the actual outcome differs from the expected one by more than four standard deviations.

Unfortunately, the standard deviation test has not brought about the mathematical certainty that its name might imply. The definition of the relevant labor market for the contested employment decision has proven to be a hotly disputed issue. Sometimes courts hold that it should be the entire workforce, sometimes it is limited to particular job categories affected by the decision at issue, but the consensus among the federal judiciary is that the relevant labor market must be defined on a case-by-case basis. In the example given above, assume that the layoffs affected only managerial and supervisory employees, eighty percent of whom are age forty or older. If the relevant labor market is defined as managerial and supervisory employees, then the inference of discrimination disappears.

As noted, courts are lenient in evaluating the intellectual rigor of the plaintiff’s methodology, readily overlooking the failure to eliminate nondiscriminatory variables like educational qualifications through multiple regression analyses. When such nondiscriminatory factors are ignored, the standard deviation test proves merely that a discrepancy from the expected outcome is unlikely to have occurred by chance and that either (a) the discrepancy is attributable to one or more factors that correlate with age and for which the analysis has not accounted, (b) the discrepancy is attributable to a combination of age and factors that correlate with age and for which the analysis has not accounted or (c) age was a factor in the employment decision. Given the inconsistency with which judges and reviewing courts have applied the standard deviation test, it is not surprising that juries have reached inconsistent results.
Moreover, the standard deviation test is by no means a required method. Rather, courts have permitted plaintiffs to introduce evidence of less scientific age comparisons.\textsuperscript{56} At least one court has suggested that the EEOC’s four-fifths rule might apply to ADEA disparate impact claims even though, by its terms, the rule applied only to Title VII claims.\textsuperscript{57} The EEOC rule presumes disparate impact whenever the selection rate for the protected group “is less than four-fifths (or eighty percent) of the rate for the group with the highest rate.”\textsuperscript{58} For example, if ninety-five percent of workers under age forty were retained during an RIF, but only seventy-five percent of workers age forty or older were retained, the RIF would be presumed to have a disparate impact on older workers.

IV. The Cumulative Effect of Multiple Employees Claiming Discrimination against the Same Employer

As damning as statistical evidence can be in an age discrimination case, a related form of evidence is potentially more prejudicial: the parade of many older workers before a jury, all claiming that the employer discriminated against them because of their age. Each plaintiff’s claim, taken individually, may have little or no merit, but it is a rare jury that can withstand the cumulative effect of hours and days spent watching an interminable procession of long-time employees all testifying that their employer is biased against older persons. No matter how thoroughly the employer justifies the employment decisions involving each plaintiff, the jury is unlikely to believe that all those older workers are imagining the same problem.

Age discrimination plaintiffs may proceed jointly against their employer as individual co-plaintiffs.\textsuperscript{59} Rule 20(a) of the Federal Rules of Civil Procedure states that multiple plaintiffs may join in a single action so long as (a) they claim relief with respect to “the same transaction, occurrence, or series of transactions or occurrences,” and (b) there are common questions of law
or fact. Courts liberally construe these requirements in favor of consolidating claims against an employer. Age discrimination plaintiffs also may seek to bring an opt-in class action on behalf of themselves and others “similarly situated.”

An employer has limited options to minimize the prejudicial impact of multiple discrimination claims being tried in a single case. It may move for separate trials under rule 42(b) or 20(b) or for severance under rule 21, but such motions are granted only occasionally. Usually age discrimination plaintiffs are allowed to try their claims together in a single case, and courts rarely acknowledge the prejudicial impact of multiple employees asserting discrimination claims against the same employer. The judge may instruct the jury to consider each plaintiff’s claim separately on its own merit, but such instructions are difficult to follow and, even with special verdict forms, impossible to monitor or enforce.

Statistical evidence can actually be the most effective way for employers to counter the cumulative effect of testimony by multiple employees alleging age discrimination. A parade of twenty-five employees might generate significant momentum in their favor at the start of the trial, but that momentum could easily turn around when the employer’s statistician shows that the company treated its 250 employees age forty or older at least as favorably as its younger employees. In this way, the employer is able to place the employees’ testimony in perspective for the jury. Statistics can be the best defense, but only if the employer has remained vigilant about the statistical impact of its decisions before litigation is filed.
V. Conclusion

Regardless of the eventual outcome of the current circuit split on the viability of age discrimination claims brought under the disparate impact theory, employers should continue to monitor the statistical impact of employment decisions, such as RIFs, on workers of various ages. While there may be no way to prevent employment litigation altogether, evaluating the statistical impact of major employment decisions before they are implemented gives the employer the opportunity to correct any significant unexplainable discrepancies. Employers will then be in a position to defend successfully against suits alleging age discrimination.
Carla J. Rozycki received her J.D. from Northwestern University School of Law in 1977. Ms. Rozycki is a partner with Jenner & Block, LLC, Chicago, Illinois, representing public and private employers in all aspects of employment law and litigation.

Patricia A. Bronte received her J.D. from Northwestern University School of Law in 1987. Ms. Bronte is a partner with Jenner & Block, LLC, Chicago, Illinois, representing public and private employers in all aspects of employment law and litigation.


6 122 S. Ct. 1290, 1290 (Apr. 1, 2002).


8 Hicks, 509 U.S. at 507; Wards Cove Packing Co., 490 U.S. at 670.


Id. at 1326.

Id. at 1325.

Id.

Id. at 1325-26.

Id. at 1326-31.

Id. at 1326; see Mullin v. Raytheon Co., 164 F.3d 696, 703-04 (1st Cir. 1999), cert. denied, 528 U.S. 811 (1999); E.E.O.C. v. Francis W. Parker School, 41 F.3d 1073, 1078 (7th Cir. 1994), cert. denied, 515 U.S. 1142 (1995); Ellis v. United Airlines, Inc., 73 F.3d 999, 1009 (10th Cir. 1996), cert. denied, 517 U.S. 1245 (1996).

Mullin, 164 F.3d at 703; Francis W. Parker School, 41 F.3d at 1077-78; Ellis, 73 F.3d at 1008.

Mullin, 164 F.3d at 703; Ellis, 73 F.3d at 1008.

Mullin, 164 F.3d at 703; Ellis, 73 F.3d at 1008.

164 F.3d at 703.

See District Council 37, Am. Fed’n of State, County & Mun. Employees v. New York City Dep’t of Parks and Recreation, 113 F.3d 347, 351 (2d Cir. 1997); Smith v. City of Des Moines, 99 F.3d 1466, 1470 (8th Cir. 1996); Mangold v. California Pub. Utilities Comm’n, 67 F.3d 1470, 1474 (9th Cir. 1995); Rose v. Wells Fargo & Co., 902 F.2d 1417, 1421 (9th Cir. 1990). Interestingly, the tie-breaking vote in Adams was cast by Judge Frank J. Magill of the U.S. Court of Appeals for the Eighth Circuit, sitting in the Eleventh Circuit by designation. Judge Magill was not a member of the panels that decided Smith or the case on which Smith relied. Houghton v. Sipco, Inc., 38 F.3d 953, 958-59 (8th Cir. 1994).

District Council 37, 113 F.3d at 351 (relying on pre-Hazen Circuit precedent); Smith, 99 F.3d at 1470 (relying on Houghton); Houghton, 38 F.3d at 958 (applying disparate impact theory to ADEA claim without discussion or analysis); Mangold, 67 F.3d at 1474 (relying on Rose, pre-Hazen Circuit precedent); Rose, 902 F.2d at 1421 (relying on Palmer v. United States, 794 F.2d 534, 538-39 (9th Cir. 1986) (no analysis)); see also EEOC v. Borden’s, Inc., 724 F.2d 1390,
1394 (9th Cir. 1984) (noting general similarities between ADEA and title VII); Geller v. Markham, 635 F.2d 1027, 1032 (2d Cir. 1980) (same), cert. denied, 451 U.S. 945 (1981).


32 Susan J. McGolrick, Supreme Court Hears Views on Whether ADEA Authorizes Disparate Impact Claims, DAILY LAB. REP., Mar. 21, 2002, at AA-1 (noting that Justices Stevens and Ginsburg had questioned the employees’ attorney about the specific facially neutral policy alleged to have had a disparate impact on older workers).

33 See, e.g., Hazen Paper, 507 U.S. at 610 (“[d]isparate treatment . . . captures the essence of what Congress sought to prohibit in the ADEA”).


39 Bazemore v. Friday, 478 U.S. 385, 400 (1986) (both individual and pattern or practice claims based on statistics); Hazelwood Sch. Dist. v. United States, 433 U.S. 299, 313 (1977) (pattern or practice claim); see also Hipp v. Liberty Nat’l Life Ins. Co., 252 F.3d 1208, 1228 (11th Cir. 2001) (noting that “statistical evidence is often extremely useful” in pattern or practice cases, but declining to hold that it is required in all such cases); Bell v. Envtl. Prot. Agency, 232 F.3d 546, 553 (7th Cir. 2000) (“[i]n a pattern and practice disparate treatment case, statistical evidence constitutes the core of a plaintiff’s prima facie case”); Adams v. Ameritech Servs., Inc., 231 F.3d 414, 423 (7th Cir. 2000) (“statistical evidence can be very useful to prove discrimination” in cases alleging either individual disparate treatment or a pattern or practice of discrimination).


41 See, e.g., Bell, 232 F.3d at 554-55 (reversing summary judgment for employer in title VII case and remanding for trial because of trial court’s rejection of employees’ statistical evidence); Adams, 231 F.3d at 428 (same in ADEA case).


45 Bazemore, 478 U.S. at 400 (holding that plaintiff’s statistical analysis need not include all variables that could explain wage differentials); Bell, 232 F.3d at 554 (reversing trial court’s summary judgment and rejection of plaintiff’s statistical evidence that considered data for positions not involved in case and showed discrepancy that was “not statistically significant”); Adams, 231 F.3d at 428 (noting that flaws in plaintiff’s expert’s methodology went to the weight but not the admissibility of the statistical evidence); Bruno v. W.B. Saunders Co., 882 F.2d 760, 767 (3d Cir. 1989) (affirming reliance on plaintiff’s statistical evidence despite failure to account for objective job qualifications); see also Alexander v. Fulton County, Georgia, 207 F.3d 1303, 1329 (11th Cir. 2000) (holding that trial court’s jury instructions ameliorated any possible prejudice to employer due to court’s erroneous admission of plaintiffs’ five pieces of irrelevant statistical evidence).

46 Bazemore, 478 U.S. at 400; but see Smith v. Xerox Corp., 196 F.3d 358, 371 (2d Cir. 1999).

47 See, e.g., Hazelwood, 433 U.S. at 313; Adams, 231 F.3d at 425; Smith, 196 F.3d at 366; Bruno, 882 F.2d at 767; Pace v. Southern Railway System, 701 F.2d 1383, 1386-87 (11th Cir. 1983), cert. denied, 464 U.S. 1018 (1983).

48 Bell, 232 F.3d at 553.

49 See Hazelwood, 433 U.S. at 309, n.14; Adams, 231 F.3d at 424; Smith, 196 F.3d at 366.

50 Smith, 196 F.3d at 365-66.

51 See Hazelwood, 433 U.S. at 313 (discussing relevant labor market extensively but remanding issue to district court for determination based on all the facts and circumstances); Adams, 231 F.3d at 425 (noting wide range of definitions of relevant labor market); Bell, 232 F.3d at 553 (holding that district court improperly limited relevant labor market to particular category of positions at issue).

52 Id.

53 Smith, 196 F.3d at 365-66.
Adams, 231 F.3d at 425 (reversing summary judgment for employer and stating, What [the plaintiff’s expert] did not do . . . was to run a multiple-regression analysis that would have isolated the relevance of age as a factor in the companies’ decisions. While this omission strikes us as odd, we are not prepared to hold as a matter of law that nothing but regression analyses can produce evidence that passes the Daubert and Kumho Tire thresholds.)

Id.; see Smith, 196 F.3d at 371 (affirming summary judgment for employer where plaintiff’s statistical analysis failed to consider other possible causes for the high percentage of layoffs among older workers).

Radue, 219 F.3d at 616 (“[b]ecause the occurrence of adverse employment actions may correlate to older employees for reasons other than intentional discrimination, causation is suggested only when the other variables are shown to be insignificant” (citation omitted)).

Stratton v. Dep’t for the Aging for the City of New York, 132 F.3d 869, 877 (2d Cir. 1997) (affirming $1.5 million jury verdict for single plaintiff who prepared age-comparison charts showing 4-1/2 year median age drop in 15 months among those listed on organization chart); Bell, 232 F.3d at 554 (holding that statistical evidence may be probative of discrimination even though the disparities are not statistically significant); but see Martinez v. Wyoming, Dep’t of Family Servs., 218 F.3d 1133, 1138-39 (10th Cir. 2000) (rejecting plaintiff’s statistical evidence based on anecdotes and data relating to wrong time frame).

Smith, 196 F.3d at 365; 29 C.F.R. § 1607.4(D) (2001) (“a selection rate for any race, sex, or ethnic group . . . ”).

29 C.F.R. § 1607.4(D).

FED. R. CIV. P. 20(a).


Coleman v. Quaker Oats Co., 232 F.3d 1271, 1297 (9th Cir. 2000) (affirming severance of ADEA claims of ten employees who worked in six states because of “potential prejudice to Quaker created by the parade of terminated employees and the possibility of factual and legal confusion on the part of the jury”), cert. denied, 533 U.S. 950 (2001); Morris v. Northrop Grumman Corp., 37 F. Supp. 2d 556, 581 (E.D.N.Y. 1999) (severing unrelated title VII claims and noting possible prejudice to employer if strong claims are tried with weaker ones); but see
Mosley, 497 F.2d at 1334 (holding that district court abused its discretion in severing claims of ten plaintiffs in race discrimination case).

63  See, e.g., Alexander, 207 F.3d at 1325-26; Duke, 928 F.2d at 1420-21.

64  Duke, 928 F.2d at 1421.