The "Substantially Younger" Requirement in O'Conner v. Consolidated Coin Caterers Corp.: Will ADEA Plaintiffs Lose Again?

Kurt Schaub

Follow this and additional works at: http://scholarlycommons.law.hofstra.edu/hlelj

Part of the Law Commons

Recommended Citation
Available at: http://scholarlycommons.law.hofstra.edu/hlelj/vol16/iss1/7

This document is brought to you for free and open access by Scholarly Commons at Hofstra Law. It has been accepted for inclusion in Hofstra Labor and Employment Law Journal by an authorized administrator of Scholarly Commons at Hofstra Law. For more information, please contact lawcls@hofstra.edu.
I. INTRODUCTION

For more than thirty years, the Age Discrimination in Employment Act ("ADEA" or "the Act") has been the law of the land, seeking to combat age bias in the workplace. Over the years, the ADEA has been expanded to cover a wider range of people. When it was first passed in 1967, the Act protected people who were between forty and sixty-five years old. The ceiling for ADEA protection was raised to seventy years of age in 1978, then removed altogether in 1986. In fact, the ADEA enjoyed widespread support from the public. However, there are now two powerful social trends in place which are going to make public support for the ADEA even stronger. The first trend is the changing age profile of the American public. The second is an economic transformation which is undermining American job stability.

Both of these changes are coming up against a counter-trend in the courts which is making the ADEA less available to plaintiffs. This Note will

2. See id. § 621(4)(b). "It is . . . the purpose of this [Act] to promote employment of older persons based on their ability rather than their age." Id.
3. See id. § 631(a).
6. See, e.g., McKennon v. Nashville Banner Publ'g Co., 513 U.S. 352, 357 (1995) (noting that the ADEA "reflects a societal condemnation of invidious bias in employment decisions"); Thornbrough v. Columbus & Greenville R.R., 760 F.2d 633, 637 (5th Cir. 1985) (noting that passage of the ADEA is a recognition that "one of the tests of a civilized society is its treatment of the elderly").
7. See infra notes 17-27 and accompanying text.
8. See infra notes 28-34 and accompanying text.
give a brief description of how demographics and economics will increase the demand for the ADEA as a remedy and then examine how a third social trend, this one legal, has been working to limit the ADEA’s availability. The trend to narrow the ADEA is evident in several recent Supreme Court decisions which have made it more difficult for age discrimination plaintiffs to make their prima facie case. Specifically, this Note will focus on the recent Supreme Court decision in O’Connor v. Consolidated Coin Caterers Corp., a case which is generally thought of as helpful for plaintiffs, and will demonstrate how it could actually make things harder for plaintiffs.

The Supreme Court’s holding in O’Connor is a bit of a paradox. One the one hand, it clearly helps plaintiffs by expanding the number of situations in which the ADEA may apply. The O’Connor decision does this by no longer requiring that plaintiffs who have been laid off, demoted, etc. in favor of younger employees show that those younger employees were outside of the protected age group (i.e., under forty). In practical terms this means that employees in their fifties and sixties who are laid off will still be able to sue under the ADEA if their replacements are also over the age of forty (as will usually be the case).

However, at the same time that O’Connor opens up the number of situations where the ADEA might apply, the decision also contains language that may harm plaintiffs. It could do this by placing limits on the statistics that plaintiffs are allowed to use when making their prima facie case. Limitations could arise because the O’Connor ruling requires that plaintiffs show that the alleged beneficiaries of employer age discrimination are employees who are “substantially younger” than the plaintiffs. To demonstrate how this requirement can affect the plaintiffs’ case, this Note will set up a hypothetical layoff situation and run an analysis on the numbers both with and without a “substantially younger” requirement.

---

9. See infra notes 72-88 and accompanying text.
10. See infra notes 72-88 and accompanying text.
12. See O’Connor, 517 U.S. at 312.
13. See McCorstin v. United States Steel Corp., 621 F.2d 749, 753-54 (5th Cir. 1980) (noting that “[s]eldom will a 60-year old be replaced by a person in [his] twenties”).
14. See O’Connor, 517 U.S. at 313. “In the age-discrimination context . . . an inference of [discrimination] can not be drawn from the replacement of one worker with another worker insubstantially younger.” Id.
15. See infra Section V.
Finally, a few policy suggestions will be made for lower courts applying *O'Connor* so that the decision can remain beneficial to plaintiffs.16 With the right interpretation of "substantially younger," the *O'Connor* requirement will not further limit the availability of the ADEA. It is important for the ADEA to be interpreted in a "plaintiff friendly" way, given the new social trends which Americans are facing. These trends are sure to increase both age discrimination and the need for the Act's protection.

II. BABY BOOMERS AND THE ADEA

The first social trend to increase demand for the ADEA is the result of changing demographics. The Baby Boomers have now aged to the point where even the youngest members of that generation will soon be within the ADEA's protected class.17 While this population bulge passes the forty year mark, medical innovation continues to increase the expected American lifespan.18 There is even evidence of an increasing average "health span" — the number of years that people remain in good health.19 This new longevity means that present-day retirees are spending more healthy years in retirement than ever before. These longer lifespans are creating similar expectations for long retirements among the Baby Boom generation.20

Some of these expectations, however, may be foiled by other Baby Boomer circumstances. A recent study found that Baby Boomers will have to triple their current rate of savings, on average, in order to maintain their present standards of living after retirement.21 But in fact, since

---

16. See infra Section VI.
August of 1997, in the midst of what has been one of the longest economic expansions since World War II, the savings rate in the U.S. has only been 3.8% of personal income. This rate is in contrast to the personal savings rate of more than 7% which Americans maintained throughout the 1960s. The result can be seen in another survey which found that 46% of all working-age Americans have saved less than $10,000 for retirement and that only 30% of 51- to 61-year olds have put aside that same amount.

As a result, barring some sudden increase in the ability to save more for retirement, Baby Boomers will be demanding longer working careers and later retirements. More and more Boomers will try to work past the age of sixty-five so they can accumulate the savings needed to cover longer retirement periods. As the federal government increases the minimum Social Security retirement age in order to keep the system solvent, the demand for more working years will only be enhanced.


23. See Easterlin et al., supra note 21, at 499-504. Since people in their forties and fifties have the highest incomes and savings rates, they are the group that is most responsible for the national savings rate. See id.


26. See Baby Boomers Not Saving, ORLANDO SENTINEL, Feb. 11, 1998, at B1. The article cites a survey done by Scudder Kemper Investments Inc., with help from Christopher Hayes of the National Center for Women and Retirement Research. See id. The survey of 1000 Boomers households with annual incomes of over $30,000 found that 64% of respondents had no idea how much money they would need to retire, 76% said that they were worried about their financial future, 68% admitted that they had planned insufficiently for retirement. See id.

27. See Roger Rosenblatt, Social Security Needs to Adapt, Fed Chief Says, Chi. Sun-Times, Nov. 21, 1997, at 36. The retirement age is scheduled to be slowly adjusted upward. See id. It will be at least 67 by the year 2027. See id. Chairman of the Federal Reserve Alan Greenspan, among others, has said that the minimum retirement age will have to be raised to above 67 in order for the
At the same time that aging Boomers seek to extend their years in the workforce, a second great social change is underway. The "downsizing" movement has come in full force, eroding job security and altering the old relationships that existed between corporate employers and American workers. Feeling constant pressure to maximize profits, employers have looked for every opportunity to cut costs through layoffs. Studies show older employees have been targeted with the greatest frequency.

The clash of demographics on the one hand, with the new economic forces on the other, has led many employment lawyers to predict that there will be an upswing in the number of ADEA lawsuits filed. System to remain solvent. See id. Senator Judd Gregg (R-NH) has recommended a retirement age of 70. See id.


29. See Kenneth A. Swinnerton & Howard Wial, Is Job Stability Declining in the U.S. Economy?, 48 INDUS. & LAB. REL. REV. 293, 303 (1996) (confirming in an empirical study that there has been greater job instability since the late 1980s).

30. See id. at 304. "[I]f the pattern [since] the late 1980s persists, workers who have stable, long-term jobs will make up an increasingly exclusive club." Id.

31. See Uchitelle & Kleinfield, supra note 28, at 26 (attributing job loss in part to Wall Street's insistence that companies elevate profits).

32. See BUREAU OF LABOR STATISTICS, 1994 Displaced Worker Survey, tbl.I, U.S. DEPT OF LABOR (Feb. 1994). Table I was republished in Gary Minda, Opportunistic Downsizing of Aging Workers: The 1990s Version of Age and Pension Discrimination in Employment, 48 HASTINGS L.J. 511, 521 (1997). Displacement rates for workers who were 55 and older, for example, were considerably higher during the recessionary period of 1991-92 than the recessionary period of 1981-82. See id. Again, this was true despite the fact that the recessionary period in 1981-82 was more severe than the 1991-2 recession. See COUNCIL OF ECONOMIC ADVISERS, supra note 28 at 6.

33. See Kirstin Downey Grimsley, Boomers' Nightmare: Age Discrimination, SALT LAKE CITY TRIB., May 25, 1997, at E1. "Companies need to focus on the fact that they may have a growing problem as the baby boom [generation] ages.... Businesses will see a growth in the number of age-discrimination cases and claims." Id. (quoting Stephen Bokat, general counsel of the National Chamber of Commerce's Litigation Center); Lorraine LaFemina, Employees, Employers Belabor Labor Laws, LI BUS. NEWS, Mar. 3, 1997, at 21. "Across the country, I see age discrimination relating to downsizing the number one complaint of workers... [because] companies have used bad judgment in methodology when firing employees." Id. (quoting Murray Port-
Some people have even speculated that the Baby Boomers’ historically high level of education (hence awareness of statutory rights), combined with a group tendency to go “against the grain” will contribute to the number of lawsuits.  

III. NARROWING THE APPLICATION OF THE ADEA

Demographics and downsizing are the two forces which will drive a growing demand for age discrimination lawsuits. However, they will both come up against a movement in the courts to restrict the application of the ADEA and discrimination cases, in general. These restrictions have not slowed the enormous publicity given to spectacular judgments and pre-trial settlements in discrimination cases, but the reality is that such suits have become more difficult to successfully prosecute. Even though the ADEA has been limited by a series of court decisions interpreting the Act, it still suffers under the same myths that continue to surround all discrimination cases — that employers are being swamped with suits which plaintiffs usually win and that even when cases settle they cost employers a fortune. The reality can be

34. See Lisa Stansky, New Age Woes: Lawyers Are Preparing Now for a Possible Wave of Age Discrimination Suits by Baby Boomers, 83 A.B.A. J. 66 (1997). “Baby Boomers have been raised on the notion of individual rights ... go to court, stand up for your rights. If cultural orientation matters, then I think we can expect that baby boomers are going to be suing a lot.” Id. at 67 (quoting Howard Eglit, professor at Chicago-Kent College of Law).

35. See infra notes 72-88 and accompanying text.

36. See Amy Saltzman, Suppose They Sue?: Why Companies Shouldn’t Fret So Much About Bias Cases, U.S. News & World Rep., Sept 22, 1997, at 68, 68-70 (finding that the media’s depiction of discrimination cases is distorted).

37. See id. at 68.

38. See id. at 69.

39. See infra notes 72-88 and accompanying text.

40. See Saltzman, supra note 36, at 69. In 1996, for example, there were less than six employment complaints filed with the EEOC for every 10,000 workers. See id. The 77,990 complaints filed that year for all types of discrimination claims (race, sexual harassment, disability, etc.) need to be considered in the context of the size of the American workforce which had reached a record 129.7 million people (11 million more than in 1992). See id.

41. See id. at 69. “[J]ob discrimination cases remain one of the single most unsuccessful classes of litigation for plaintiffs. They settle less and lose more than almost anything else.” Id. (quoting Theodore Eisenberg, Cornell University law school professor).

42. See Saltzman, supra note 36, at 69. James Dertouzos, a senior economist at the Rand Institute, conducted a survey of 470 wrongful dismissal cases in California between 1987 and 1994. See id. The “wrongfulness” alleged included all forms of bias, not just age discrimination. See id. Dertouzos found that 17% of cases were dropped after initial complaint, costing employers
seen in surveys which find an increasing ratio of cases brought to cases won.\textsuperscript{43} Even the number of class action suits filed by the EEOC has plummeted.\textsuperscript{44}

To understand how the courts have limited the use of the ADEA in the past few years, we must first look at how the burden of proof for ADEA cases was set up in the past and how recent court decisions have altered it. Traditionally, there have been three ways of establishing an ADEA claim:\textsuperscript{45} (1) through direct evidence of discriminatory intent which had an adverse effect on the plaintiff (i.e., a “smoking gun”); (2) through the use of circumstantial evidence to show intent;\textsuperscript{46} and (3) through statistical proof of a policy or employer action which had a disparate effect (though not necessarily intentionally) on the protected group.\textsuperscript{47} Methods (1) and (2) are known as “disparate treatment,” and method (3) as “disparate impact.”\textsuperscript{48} Method (1) usually relies on “stray remarks” from the plaintiff’s employer which show that the layoff is due to the plaintiff’s age.\textsuperscript{49} These cases are becoming more rare as em-
Because employers are becoming more cautious, the most common type of ADEA case uses circumstantial evidence [method (2)] to make a case of disparate treatment. It typically involves the use of statistics to draw an inference of employer discrimination. The structure of proof for circumstantial "disparate treatment" cases under the ADEA has followed the same disparate treatment framework laid out in Title VII cases. The basic framework was first outlined in the Supreme Court case of *McDonnell Douglas v. Green*.

**A. McDonnell Douglas v. Green: The Basic Framework**

The ADEA was passed only three years after the passage of the 1964 Civil Rights Act which included Title VII prohibitions against discrimination based on "race, color, religion, or national origin." Ever since the passage of the two statutes, the courts have noticed that the ADEA and Title VII are designed to ban discrimination in ways which parallel each other. As a result, for many years the proof scheme required for the ADEA has imitated the one required by Title VII, the so-called "McDonnell Douglas test."

In *McDonnell Douglas v. Green*, the Supreme Court set up a three-step burden-shifting operation for proving cases under Title VII. In the first step, the burden is on the plaintiff to make a prima facie case of

50. See id.
53. See id.
54. See Lorillard v. Pons, 434 U.S. 575, 584 (1978) (stating that "the prohibitions of the ADEA were derived in haec verba from Title VII."); see also Oscar Meyer & Co. v. Evans, 441 U.S. 750, 756 (1979) (observing that the ADEA and Title VII "share a common purpose, the elimination of discrimination in the workplace").
55. See, e.g., Pena v. Brattleboro Retreat, 702 F.2d 322, 323-324 (2d Cir. 1983); Smith v. Flax, 618 F.2d 1062, 1066 n.3 (4th Cir. 1980); Loeb v. Textron, Inc., 600 F.2d 1003, 1008 (1st Cir. 1979); Cova v. Coca-Cola Bottling Co., 574 F.2d 958, 959 (8th Cir. 1978); Rodriguez v. Taylor, 569 F.2d 1231, 1239 (3d Cir. 1977).
56. See McDonnell Douglas v. Green, 411 U.S. 792 (1973) (giving the basic framework for a Title VII case involving circumstantial evidence); see also 45C AM. JUR. 2D at § 2738 (explaining how courts have borrowed the McDonnell Douglas framework to use in ADEA cases). A case for intentional discrimination can also be made with direct evidence in which case the whole apparatus of *McDonnell Douglas* can be dispensed with. See 45C AM. JUR. 2D at § 2739. It is rare, however, for ADEA cases to rely primarily on direct evidence. See Posner, supra note 49, at 335.
57. See McDonnell Douglas, 411 U.S. at 802.
discrimination. McDonnell Douglas was a racial discrimination case, so the Court ruled that the plaintiff had to show that: (1) he belonged to a racial minority, (2) he was qualified for the job he applied for, (3) despite his qualifications, he was rejected for the job, and (4) after the rejection, the position remained open and the employer continued to seek applicants from persons with the plaintiff's qualifications.

In an age discrimination case, the McDonnell Douglas framework is slightly modified. In that case, the plaintiffs claiming disparate treatment must make a prima facie case by showing that they: (1) were forty or older, (2) were qualified to do the job, (3) were fired (or not hired, promoted, etc.) and (4) were replaced by people outside of the protected class. In O'Connor, the Supreme Court ruled that element (4) did not apply to the ADEA. This means that in the future, plaintiffs will not have to show that replacement workers came from outside of the protected class.

In general, the amount of proof that the plaintiff must present to make the prima facie case is considered minimal. The prima facie case has been kept as a low hurdle because the courts realize that it is rare for the plaintiff to have a "smoking gun" — direct evidence of discriminatory intent. If the plaintiff can show the four elements listed above and explain how they are logically connected to an inference of discrimination, then they have made a prima facie case. After the plaintiff does this, the McDonnell Douglas scheme shifts the burden of production to

58. See id.
59. See id.
60. See 45C AM. JUR. 2D at § 2738.
61. See O'Connor v. Consolidated Coin Caterers Corp., 517 U.S. 307 (1997). O'Connor changed this branch of the prima facie case so that it only requires that the replacement was "substantially younger." See id. at 312. To simplify the discussion of the ADEA, the remainder of this Note will focus on ADEA plaintiffs who bring suit for unlawful discharge because that was the factual case in O'Connor. That does not mean, of course, that plaintiffs who bring suit because they were not hired, promoted, etc. are not similarly affected by O'Connor.
62. See id. at 312. In O'Connor, the plaintiff was 56-years old and the alleged beneficiary of the discriminatory hiring was 40 years old. See id. at 309-10. The replacement worker was a member of the ADEA's protected class, but that did not mean that plaintiff had failed to make a prima facie case. See id. at 312.
63. See Diaz v. American Tel. & Tel., 752 F.2d 1356, 1361 (9th Cir. 1985) ("establishing a prima facie is not designed to be onerous"). The plaintiff only needs to produce evidence that suggests that the employment action was done for discriminatory reasons. See id.
64. See, e.g., United States Postal Serv. Bd. of Govers. v. Aikens, 460 U.S. 711, 716 (1983) (commenting that "[t]here will seldom be 'eyewitness' testimony as to the employer's mental processes.").
65. See Texas Dep't of Community Affairs v. Burdine, 450 U.S. 248, 254 n.7 (1981) (producing enough evidence to permit inference of discrimination is plaintiff's burden).
the defendant. A rebuttable presumption is created that the employer's actions were not allowable, unless the employer can show that they were taken for legitimate, non-discriminatory reasons.

If the defendant successfully rebuts the presumption that its actions were been discriminatory, the third branch of McDonnell Douglas applies and the burden of production is shifted back to the plaintiff. Originally, if the plaintiff could show that the explanations offered by the defendant were merely "pretextual," then plaintiff was entitled to judgment as a matter of law. Showing that the defendant's explanations were a mere "pretext" did not mean that the plaintiff must show that the actual reasons were discriminatory, only that the reasons offered by the defendant were not genuinely responsible for the decision. This holding was changed by the Supreme Court's decision in St. Mary's Honor Center v. Hicks.

B. St. Mary's Honor Center v. Hicks: "Pretext Plus"

In St. Mary's Honor Center v. Hicks, the Supreme Court decided it was not enough for the plaintiff to show that the defendant's offered explanations were "pretextual." Plaintiff had to offer proof that the actual motive was discriminatory and that "but for" the discriminatory actions taken by the employer, the employee would not have been laid off.

In Hicks, a black employee was fired from his job at a halfway house and the proffered reason was his violation of institutional rules. More serious rule violations by other employees, however, had been ignored. The Eighth Circuit ruled that it was enough for plaintiff to show that the reason proffered by the defendant was not the real reason; the

67. See id.
68. See id.
69. See Hicks v. St. Mary's Honor Ctr., 970 F.2d 487, 492 (8th Cir. 1992) (finding that "once plaintiff proved all of defendant's proffered reasons for the adverse employment actions to be pretextual, plaintiff was entitled to judgment as a matter of law"). The Supreme Court overturned the Eighth Circuit in St. Mary's Honor Center v. Hicks, 509 U.S. 502, 511 (1993) (holding that plaintiff was not entitled to judgment as a matter of law).
70. See Hicks, 970 F.2d at 492.
71. 509 U.S. 502 (1993); see infra notes 72-79 and accompanying discussion.
72. See Hicks, 509 U.S. at 507.
73. See id. at 508.
75. See Hicks, 509 U.S. at 508.
defendant's rebuttal failed and plaintiff was given verdict as a matter of law. The Supreme Court, however, ruled that even where the plaintiff has shown that the defendant's explanation was a mere pretext, the plaintiff must still show, by a preponderance of the evidence, that discrimination was the actual reason the plaintiff was fired.

As Justice Souter pointed out in his dissent, the Supreme Court majority in *Hicks* "adopt[ed] a scheme . . . unfair to plaintiffs, unworkable in practice, and inexplicable in forgiving employers who present false evidence in court." Requiring this "pretext plus" has made the plaintiffs' case under ADEA particularly difficult because the "actual" reason may be discriminatory, but at the same time highly entangled with non-discriminatory reasons. Making the case for plaintiffs was also made more difficult, however, by another Supreme Court decision made the same year — *Hazen Paper Co. v. Biggins*.

**C. Hazen Paper Co. v. Biggins: Permitting Close Proxy to Age**

In *Hazen Paper Co. v. Biggins*, a company fired a sixty-two year old employee who was just a few weeks away from vesting in the company's pension plan. The company pension plan required employees to have ten years of experience with the company before the plan would vest. A unanimous Court ruled that since the pension plan vested for all employees with sufficient experience, not age, there was nothing in the company's behavior that was actionable under the ADEA. The Court reasoned that "[b]ecause age and years of service are analytically distinct, an employer can take account of one while ignoring the other, and thus it is incorrect to say that a decision based on years of service is necessarily age-based."

---

76. See *Hicks*, 970 F.2d at 492.
77. See *Hicks*, 509 U.S. at 511.
78. Id. at 533.
80. See *Hazen*, 507 U.S. at 607.
81. See id.
82. See id. at 611. The Court decided that firing the employee to prevent pension plan from vesting was actionable, however, under the Employee Retirement Income Security Act (ERISA). See id. at 612.
83. *Hazen*, 507 U.S. at 611 (emphasis added).
To summarize, after *Hicks* and *Hazen*, an ADEA plaintiff must make a prima facie case for discrimination. This prima facie case creates a rebuttable presumption that the defendant acted in an illegal manner. If the employer offers a legitimate reason for its actions and the plaintiff can prove that the employer’s legitimate reason could not have been the genuine reason, the burden of production, nevertheless, shifts back to the plaintiff. For example, if the employer says that plaintiff was laid off to save costs, cost-saving will be a legitimate reason. After *Hazen*, though, even if the plaintiff can show she has been replaced by a younger employee who earns ninety-nine percent of the plaintiff’s old salary, the employer has still shifted the burden of production back to the plaintiff. The strong circumstantial evidence that defendant was doing more than cutting costs will not stop the burden of production from being shifted back to plaintiff. The plaintiff must then dig deeper for the real reason for the lay-off and show that “but for” discrimination, he would not have been fired. Plaintiff’s burden is especially hard since pay level and age will probably have a strong correlation. Even without direct evidence of the employer’s mindset, the plaintiff must separate out the effect of the two variables (pay level and age) and show that age was a “but for” cause. The end result is that *Hazen* has restricted the “range of circumstantial evidence upon which a factfinder can draw the inference of discrimination.”

**IV. O’CONNOR V. CONSOLIDATED COIN CATERERS CORP. AND THE PLAINTIFFS’ PRIMA FACIE CASE**

In the midst of all of these Supreme Court cases (*Hicks*, *Hazen*, etc.) that have made the plaintiffs’ task more difficult, many felt that *O’Connor* offered plaintiffs some relief. By ruling that the plaintiff does not have to show replacement by someone too young to be protected by the ADEA, *O’Connor* expanded the number of situations in which the ADEA will apply. That was much of the holding was clearly helpful for

---

85. *See Hicks*, 509 U.S. at 507. The burden of persuasion always remains with the plaintiff. *See id.*
86. *See Hicks*, 509 U.S. at 508.
87. *See id.*
89. *See O’Connor*, 517 U.S. at 312.
plaintiffs. It also followed common sense: If a forty-nine year old plaintiff is replaced by a thirty-nine year old, why should that ten-year gap be any less meaningful than the ten-year gap between a fifty-year old plaintiff replaced by a forty-year old?90 The ten year gap should be equally probative in both cases. With the O'Connor ruling, the Supreme Court emphasized that the important protection offered by the ADEA was given to individuals who were discriminated against because of their age individually, not to a particular age group (i.e., those over forty).91 Although the Court does not state it explicitly, it follows from the logic of O'Connor that plaintiffs may "subgroup" when making their case.92 The issue of subgrouping was raised in the period before O'Connor in the case of Lowe v. Commack School District.93 This case is examined in greater depth to illustrate why subgrouping could be controversial.

A. Subgrouping Older Workers in Lowe v. Commack School District

The issue of subgrouping was raised in the period before O'Connor.94 The United States Circuit Courts were divided over whether the practice of subgrouping should be allowed, but a substantial number were in favor of it.95 The Second Circuit decision in Lowe v. Commack School District is a good example of this and was later cited.

90. See Oncale v. Sundowner Offshore Servs., 118 S. Ct. 998, 1001 (1998). (holding that sexual harassment laws make it illegal for a member of one sex to sexually harass another person of the same sex). The logic used in this case parallels the thinking in O'Connor because both cases ignore the category or status in which the perpetrator of harassment (in Oncale) or beneficiary of discrimination (in O'Connor) belong to. See id. at 1001-02. "[N]othing in Title VII necessarily bars a claim of discrimination 'because of . . . sex' merely because the plaintiff and the defendant . . . are of the same sex." Id. "The fact that one person in the protected class has lost out to another person in the protected class is thus irrelevant, so long as he has lost out because of his age." O'Connor, 517 U.S. at 312.

91. See O'Connor, 517 U.S. at 312.

92. "Subgrouping" is the way this practice was described in the case of Lowe v. Commack Sch. Dist., 886 F.2d 1364, 1373 (2d Cir. 1989).

93. 886 F.2d 1364 (2d Cir. 1989) (allowing subgrouping under theory of disparate treatment).

94. See Lowe; 886 F.2d at 1364; see also Goldstein v. Manhattan Indus., Inc., 758 F.2d 1435 (allowing subgrouping under theory of disparate treatment); Douglas v. Anderson, 656 F.2d 528 (9th Cir. 1981) (allowing subgrouping under theory of disparate treatment); McCorstin v. United States Steel Corp., 621 F.2d 749 (5th Cir. 1980) (allowing subgrouping under theory of disparate treatment). But cf. Graffam v. Scott Paper, 60 F.3d 809 (1995) (allowing subgrouping under both theories of disparate impact and disparate treatment).

95. See supra note 94.
by other circuits as well. In Lowe, the Second Circuit allowed employees to make a circumstantial case for disparate treatment by comparing job applicants fifty and older to applicants who were younger. The Second Circuit's concerns in Lowe should be looked at in more detail.

In Lowe, a pair of part-time, substitute teachers tried to have their status upgraded to full-time, regular teachers. When they failed to do so, they sued their school district over the method that it had used to determine whom it would promote. The plaintiffs' case relied on both disparate treatment and disparate impact theories. At the time they applied for an upgrade in their job status, both plaintiffs were in their early fifties and both claimed that the District's hiring practices were age discriminatory. However, because many of the teachers who were actually promoted were in their forties, the plaintiffs did not allege that the district discriminated against the entire class of people protected by the ADEA, only those over fifty.

The Second Circuit decided that under a disparate impact theory, plaintiffs should only be allowed to show that an employment practice had a negative impact on the "class" explicitly protected by the ADEA — those forty and over. Because they did not do this, the plaintiffs failed to make their prima facie case. The Second Circuit said that "[u]nder [the plaintiffs'] approach... any plaintiff can take his or her own age as the lower end of a 'sub-protected group' and argue that said 'sub-group' was disparately impacted." The Second Circuit left subgrouping available as a strategy for future plaintiffs who might want to claim disparate treatment. Since Lowe, there have been other

97. See Lowe, 886 F.2d at 1372.
98. See id at 1366.
99. See id. at 1368.
100. See id. at 1370. The two theories do not exclude one another. See id.
101. See Lowe, 886 F.2d at 1367-69. These practices consisted of written tests and interviews. See id. at 1367.
102. See id. at 1372. The plaintiffs alleged that older applicants with strong performances on the written tests and interviews were rejected, while younger applicants with poorer scores were given promotions. See id. at 1368-69.
103. See id. at 1373.
104. See Lowe, 886 F.2d at 1374.
105. Id. at 1373.
106. See id. at 1374. "We reiterate that nothing we say here should be taken to preclude a plaintiff from prevailing on a disparate treatment claim where the 'beneficiaries' of the discrimination, although younger than the plaintiff, are nevertheless themselves within the protected group of those 40 and over." Id. The plaintiffs in Lowe only failed under a disparate treatment theory because their evidence was insufficient. See id. at 1377.

http://scholarlycommons.law.hofstra.edu/hlelj/vol16/iss1/7
courts which have ruled out disparate impact cases as well, and some of
them have expressed the same reservations about subgrouping as the
Second Circuit did in Lowe.\footnote{See e.g., Barnes v. Gencorp, Inc., 896 F.2d 1457, 1467 (6th Cir. 1990).}

Judge Pierce, who wrote the concurring opinion in Lowe, made a
practical argument, however, for also allowing ADEA plaintiffs to make
their claims under a theory of disparate impact.\footnote{See Lowe, 886 F.2d at 1379-80.} Judge Pierce argued that

the likely beneficiary of discrimination against a 60-year old person
will be another member of the protected group, \textit{i.e.}, a person more
than 40 years of age. Thus, "[i]f no intra-age group protection were
provided by the ADEA, it would be of virtually no use to persons in
the upper ages of the protected class whose jobs require experience
since even an employer with clear anti-age animus would rarely re-
place them with someone under 40."\footnote{Id. at 1379 (quoting Maxfield v. Sinclair Int'l, 766 F.2d 788, 792 (3d Cir. 1985)).}

The ADEA expressly protects \textit{individuals} against age discrimination
[cites omitted]... Nothing in the statute suggests that when a member
of the protected class is discriminated against on the basis of age, the
extent of that individual member's rights should be contingent upon
the age of the person who has benefited from that discrimination. The
majority opinion, however, would raise this fortuity to the level of a
dispositive factor in those instances where disparate treatment cannot
be shown... In my view, such a refusal goes against the 'thrust' of
Congress' intent, as the Supreme Court discerned that intent in
Griggs.\footnote{Id. at 1379-80. See id. at 1380. Judge Pierce cited several court cases which upheld subgroup protection
under disparate treatment and argued that the same reasoning should be applied to subgrouping
under disparate impact. See id. at 1380. Among the cases cited were Goldstein v. Manhattan Indus., Inc., 758 F.2d 1435 (11th Cir. 1985); McCorstin v. United States Steel Corp., 621 F.2d 749
73 U. Det. Mercy L. Rev. 591 (1996).}

The Supreme Court in O'Connor shared some of the same con-
cerns as Judge Pierce, but did not go so far as endorsing disparate im-
 pact in ADEA cases. They did allow for subgrouping, however, to help
make the case under a theory of disparate treatment.
B. The Effect of O'Connor v. Consolidated Coin Caterers Corp. on Subgrouping

In O'Connor, when a fifty-six year old sales manager was fired and replaced by a forty-year old, he brought an ADEA case against his employer. The sales manager's employer, Consolidated Coin, did not dispute plaintiff O'Connor's use of the McDonnell Douglas scheme described above. The only issue that the Court decided in O'Connor was whether or not the plaintiff made his prima facie case despite failing to allege that he was replaced by a person who was a non-member of the protected class. The Court decided that the ADEA was not set up to protect people in one "age class" from discriminatory treatment which benefits another outside that class. "The fact that one person in the protected class has lost out to another person outside the protected class is thus irrelevant, so long as he lost out because of his age." But the Court noted that the ADEA did limit this "age protection" (not "class protection") to people forty and over who have been discriminated against in favor of someone "not insubstantially younger."

C. What Does O'Connor Mean by "Substantially Younger"?

In its O'Connor decision, the Supreme Court ruled that intra-group comparisons could be made between members who were close in age, but that as a practical matter the prima facie case would require evidence that the replacement worker was "not . . . insignificantly younger" than the plaintiff for it to carry any weight. That raises the question of what "significantly younger" means. The Supreme Court is not specific but does offer hints. The plaintiff in O'Connor was sixteen years older than his replacement. That sixteen-year discrepancy was enough for the plaintiff's prima facie case, but the Court does not suggest that the gap must be that large.

111. See O'Connor v. Consolidated Coin Caterers Corp., 517 U.S. 307, 309 (1997). The Supreme Court noted that they had never explicitly ruled that McDonnell Douglas could be applied to ADEA cases but left the issue unresolved for now. See id.
112. See id. at 312.
113. See id.
114. See id. at 313.
115. See id. at 312.
116. See id.
117. See id.
Other examples of "significant" age differences can be found in pre- and post-O'Connor ADEA cases. In 1981, the Ninth Circuit ruled in Douglas v. Anderson\(^{118}\) that a five-year gap was enough to help make a prima facie case when a fifty-four year old worker was discharged and replaced by a forty-nine year old.\(^{119}\) Meanwhile, the Third Circuit and the D.C. Circuit split over whether eight years was enough of a difference, with the Third Circuit saying that it was in Barber v. CSX Distribution Services,\(^{120}\) and the D.C. Circuit saying it was not in Adkins v. Safeway, Inc.\(^{121}\) The court in Barber said "[t]here is no magical formula to measure a particular age gap and determine if it is sufficiently wide to give rise to an inference of discrimination, however, case law assists our inquiry."\(^{122}\) It cited the earlier case of Healy v. New York Life Insurance Co.\(^{123}\) in which nine years was considered enough to make the prima facie case even though the benefited employee was also over forty.\(^{124}\) A New York district court ruled that twelve years was "substantially younger" in McNulty v. New York City Department of Finance.\(^{125}\) However, the Eighth Circuit ruled that five years younger was not substantially younger in Schiltz v. Burlington Northern Railroad.\(^{126}\) In Hartley v. Wisconsin Bell,\(^{127}\) the Seventh Circuit decided that a ten-year difference in ages was "substantial" under O'Connor and that six or seven years might not be enough depending on the rest of the evidence.\(^{128}\)

There have also been a few unusual rulings in which the courts have allowed (or said that they would allow) prima facie evidence showing that plaintiff was younger than the person supposedly benefiting from the discriminatory conduct. The Tenth Circuit has in fact done this in Greene v. Safeway Stores\(^{129}\) and the Ninth Circuit has said that it would in dicta in the Douglas case.\(^{130}\) These courts may not be trying to...

---

\(^{118}\) 656 F.2d 528 (9th Cir. 1981).

\(^{119}\) See Douglas, 656 F.2d at 533.

\(^{120}\) 68 F.3d 694, 699 (3d Cir. 1995).

\(^{121}\) 985 F.2d 1101, 1104 (D.C. Cir. 1993).

\(^{122}\) Barber, 68 F.3d at 699.

\(^{123}\) 860 F.2d 1209 (3d Cir. 1988); Barber, 68 F.3d at 699.

\(^{124}\) See Healy, 860 F.2d at 1214.


\(^{126}\) 115 F.3d 1407, 1413 (8th Cir. 1997).

\(^{127}\) 124 F.3d 887 (7th Cir. 1997).

\(^{128}\) See Hartley, 124 F.3d at 893.

\(^{129}\) 98 F.3d 554, 557 (10th Cir. 1996) (reversing a grant of summary judgment to defendant even though the replacement worker was five years older than the plaintiff).

\(^{130}\) See Douglas, 656 F.2d at 533; Loeb v. Textron, 600 F.2d 1003, 1013 n.9 (1st Cir. 1979). "Replacement by someone older would suggest no age discrimination but would not disprove it conclusively. The older replacement could have been hired, for example, to ward off a threatened
lay the groundwork for allowing reverse discrimination under the ADEA. It is more likely that they want plaintiffs to be able to counter the strategic rehiring of laid-off employees.31

This Note will now set up a hypothetical “downsizing” situation. This hypothetical will allow us to look at how subgroup analysis can affect the statistics used to make the plaintiff’s prima facie case. The hypothetical will also examine how the “substantially younger” requirement in O’Connor could affect the plaintiff’s statistical case by requiring data to be omitted from the analysis.

V. A “DOWNSIZING” HYPOTHETICAL: HOW O’CONNOR AFFECTS THE NUMBERS IN AN AGE DISCRIMINATION CASE

Say that a large company, the Downsize Corporation, has announced that due to economic pressures, it will be forced to shrink its workforce. Imagine that the layoff will affect many thousands of people, but that within the company there is a group of 100 employees who share the same job category. They have the same responsibilities, duties, and work hours. In addition, they are paid roughly the same amount of money and have comparable amounts of experience and skills. They range in age from thirty-five to fifty-five and are evenly spread out across this range. Suppose that the layoffs produce the results shown in Table IA. Afterwards, a group of plaintiffs bring an ADEA suit under a joint theory of disparate impact and disparate treatment. They use statistics to present circumstantial evidence that they have been discriminated against.

\[\text{discrimination suit.} \quad \text{Id.} \]

31. See Loeb, 600 F.2d at 1013 n.9.
TABLE IA. Layoffs experienced by Downsize Corporation’s employees who fall within the same job category and have the same level of skill, salary, and experience.

<table>
<thead>
<tr>
<th>Employee Age Bands</th>
<th>Number of Employees in Age Band</th>
<th>Number of Employees Fired</th>
<th>Number of Employees Still Employed</th>
</tr>
</thead>
<tbody>
<tr>
<td>35-40</td>
<td>25</td>
<td>2</td>
<td>23</td>
</tr>
<tr>
<td>40-45</td>
<td>25</td>
<td>2</td>
<td>23</td>
</tr>
<tr>
<td>45-50</td>
<td>25</td>
<td>6</td>
<td>19</td>
</tr>
<tr>
<td>50-55</td>
<td>25</td>
<td>10</td>
<td>15</td>
</tr>
</tbody>
</table>

A. O'Connor Justifies the Subgrouping of Plaintiffs

Statistics can test these numbers and see if there was a connection between the employees’ ages and whether or not they lost their job. In statistical language, a statistical test should look to see if the employees’ age and job status were treated as “independent variables.” If the two variables are independent of each other, then the odds of the employee falling into a particular age category had no effect on the employees’ odds of being fired or retained.

A chi-square test can be used to test the independence of the two variables for all of the employees. The chi-square takes advantage of a law from the field of probability. If two variables, such as age and job status, are independent, then the odds that an employee is in a particular

---

132. In order to sort the employees by age, some lines need to be drawn. For the puppies of these tables, an employee who is exactly 40 is in the 35-40 category, but someone who is 40 years and a day goes in the 40-45 category.

133. The employees in Table IA can be categorized according to two different variables – age and job status. Each employee has a “job variable” (fired or not fired) and an “age variable” (35-40, 40-45, 45-50, or 50-55).

134. See Henry Garrett, Elementary Statistics 139 (Van Rees Press, New York, 1962) (giving an explanation of the chi-square statistic); Michael Piette and Douglas Sauer, Legal and Statistical Approaches to Analyzing Allegations of Employment Discrimination, 3 J. LEGAL ECON. 1 (1993). “A common first [test] in establishing a prima facie case is to use the chi-square statistic to compare two or more distributions.” Id. at 6; Richard Singleton, Use of Statistics in Age Discrimination Litigation, 371 P.L.I. Lit. 177 (1989). “The first step in a statistical analysis of a layoff generally will be to perform a two-by-two chi-square analysis, comparing numbers of employees under and over age forty laid off and not laid off. . . . If the result is statistically significant, the possibility of age discrimination is raised.” Id. at 184.
age group and a particular job category is simple to calculate — simply multiply the odds of each outcome occurring by itself.

For example, suppose that the Downsize Corp. acted legally and that age had no effect influence on whether an employee was fired. Statistically this would mean that the two variables of age and job status were independent of each other. It would also mean that the odds of an employee belonging to a particular age group and a particular job category could be calculated by multiplying the odds of those things happening separately. Suppose that plaintiffs' counsel wants to know the odds of a Downsize Corp. employee being laid-off and being between the ages of thirty-five and forty. Table IA shows that the odds of being laid-off for all employees were 20/100. The odds of an employee being between thirty-five and forty-years of age were 25/100. Take the odds of being fired, multiply by the odds of being thirty-five to forty, and the result is the odds of being both: 1/20. There were 100 people in the original workforce. Since the odds are one out of twenty will be “fired” and “thirty-five to forty,” there should be five people in that category. Similar calculations can be done for all eight of the categories. Table IB shows the same information as Table IA, but with the expected values for each cell placed in parentheses next to the number which was actually observed.

TABLE IB. Employees fired and employees retained. The expected number of employees in each category is placed in parentheses next to the actual number of employees who were in that category.

<table>
<thead>
<tr>
<th>Employee Age Bands</th>
<th>Number of Employees in Age Band</th>
<th>Number of Employees Fired</th>
<th>Number of Employees Still Employed</th>
</tr>
</thead>
<tbody>
<tr>
<td>35-40</td>
<td>25</td>
<td>2 (5)</td>
<td>23 (20)</td>
</tr>
<tr>
<td>40-45</td>
<td>25</td>
<td>2 (5)</td>
<td>23 (20)</td>
</tr>
<tr>
<td>45-50</td>
<td>25</td>
<td>6 (5)</td>
<td>19 (20)</td>
</tr>
<tr>
<td>50-55</td>
<td>25</td>
<td>10 (5)</td>
<td>15 (20)</td>
</tr>
</tbody>
</table>

135. These are the odds because 20% of the workforce lost its job.
136. These are the odds because 25% of the workforce is in that age category.
137. There are eight categories in this hypothetical because there are four age groups and two job outcomes.
In this case, the number of employees laid off in the younger age groups is less than expected and the number laid off in the older categories is more than expected to see. It looks like the two variables were not independent of each other and that employees who were older were more likely to lose their jobs. This appearance, however, is subjective.

A statistical test such as the chi-square can provide more objective evidence that the two variables are not independent of each other. The test does this by comparing results such as the ones seen in Table IB to all of the theoretical outcomes that could happen at random when the two variables are independent. If the odds of seeing results like those above will very rarely happen when the variables are independent, then courts may draw the inference that the Downsize Corp. acted illegally.

The chi-square statistic is calculated by taking the difference between the expected number for a cell and the actual number observed, squaring the difference, dividing what was expected for that cell and adding the results for all of the cells. In this case, the chi-square statistic is equal to 11.0. This value is compared to all possible values that might arise when layoff decisions are not affected by age. In this case, a table of chi-square statistics shows that the odds of seeing the Table IB pattern of layoffs when the variables are independent are between 1 in 20 and 1 in 100. If it was that unlikely to see the Table IB

138. See WAYNE C. CURTIS, STATISTICAL CONCEPTS FOR ATTORNEYS: A REFERENCE GUIDE 144-51 (1983) (providing another example of the use of chi-square to test the independence of two variables).

139. See id. at 142.

140. The chi-square statistic for Tables IB is

\[
\begin{align*}
(2 - 5)^2/5 + (2 - 5)^2/5 + (6 - 5)^2/5 + (10 - 5)^2/5 + (23 - 20)^2/20 + (23 - 20)^2/20 + (19 - 20)^2/20 + (15 - 20)^2/20 = 1.8 + 1.8 + 0.2 + 5.0 + 0.45 + 0.45 + 0.05 + 1.25 = 11.0.
\end{align*}
\]

141. The most unusual outcomes that will occur at random have large differences between the numbers that are observed and the numbers that are expected. Because the chi-square statistic is calculated by taking the difference of these two numbers, see CURTIS, supra note 138, at 142, unlikely outcomes will generate large chi-square statistics.

142. The chi-square statistic that is generated here has to be compared to a chi-square number in a theoretical distribution. This distribution has a different shape depending on the "degrees of freedom" of the distribution. See GARRETT, supra note 134, at 142. The degrees of freedom ("df") is the number of age categories minus one (a − 1) times the number of employment categories minus one (e − 1). See id. In Table IB, for example, df = (a − 1) X (e − 1) = 3 X 1 = 3. For a chi-square distribution with three degrees of freedom, there is a 5% chance that the chi-square statistic will be above 7.81 and a 1% chance that it will be above 11.34. See CURTIS, supra note 138, at 221. Since the chi-square statistic we calculated is 11.0, the probability of seeing the Table IB outcome if age and employment are unrelated is between 1% and 5%. See id. There is no specific level of probability that the courts require in order to decide that there has been discrimination, but a 5% level has been used frequently. See Richard Mariani & Kimberly Robertson, Age Discrimination Litigation: RIFs, Statistics and Stray Remarks, 64 DEF. COUNS. J. 88, 92 (1997).
outcome when the Downsize Corp. acted fairly, then that should allow a court to draw an inference that the company behaved illegally. The inference should be strong enough for plaintiffs to make a prima facie case. The categories in Table IB which were disproportionately affected are obvious. There are an increasing number of layoffs for the older categories and the trend is “one-way.”

Now look at what a pre-O'Connor court might have forced plaintiffs to do when making their case. Table II combines the downsizings from Tables IB into only two age categories: for employees under forty and those forty and over.

**TABLE II.** Comparing employees under forty to employees forty and over. The expected value for each category is in parentheses.

<table>
<thead>
<tr>
<th>Employee Age Bands</th>
<th>Number of Employees in Age Band</th>
<th>Number of Employees Fired</th>
<th>Number of Employees Still Employed</th>
</tr>
</thead>
<tbody>
<tr>
<td>under 40</td>
<td>25</td>
<td>2 (5)</td>
<td>23 (20)</td>
</tr>
<tr>
<td>40 and over</td>
<td>75</td>
<td>18 (15)</td>
<td>57 (60)</td>
</tr>
</tbody>
</table>

The chi-square statistic must be slightly altered when dealing with a 2X2 contingency table, but it will still work. In this case the chi-square statistic will be $2.08^{145}$ which is not significant at even the $p = \ldots$

---

143. This is really the 35-40 band from Table IB. The hypothetical assumes that there are no employees under 35 in this job category.

144. See Garrett, supra note 134, at 145. The 2X2 refers to the fact that there are two possible job outcomes and two possible age classifications: “fired”/“still employed” and “<40/40+.” The statistic must now be calculated from a chi-square distribution which only has 1 degree of freedom ($df = 1$). See id. at 146.

145. With only two age categories, the chi-square statistic must use a “Yates correction fac-
0.05 level. That means that if a pre-\textit{O'Connor} court refused to allow plaintiffs from looking at members of the over-forty group as different from each other, plaintiffs might not be able to make a prima facie case using statistics.

Now look at what might happen when courts accept subgrouping. The plaintiffs won't have to draw a line between employees under forty and those who are forty and over. If the plaintiffs think that the company's discriminatory treatment only affects employees who are forty-five and older, then they will be free to run a test which looks at the independence of the age categories of "35-45" / "45+" and the employment category "fired" / "still employed." This result is shown in Table III.

\textbf{TABLE III.} Comparing employees under forty-five to those who are forty-five and over. The expected value for each category is in parentheses.

<table>
<thead>
<tr>
<th>Employee Age Band</th>
<th>Number of Employees in Age Band</th>
<th>Number of Employees Fired</th>
<th>Number of Employees Still Employed</th>
</tr>
</thead>
<tbody>
<tr>
<td>under 45</td>
<td>50</td>
<td>4 (10)</td>
<td>46 (40)</td>
</tr>
<tr>
<td>45 and over</td>
<td>50</td>
<td>16 (10)</td>
<td>34 (40)</td>
</tr>
</tbody>
</table>

The chi-statistic here is 7.56 which is very significant. It is a strong indication that the company has engaged in discrimination because there is less than a one percent chance that such an outcome would happen if the company acted in an age-neutral way.\textsuperscript{147} Table IV looks at the re-

\textsuperscript{146} This means that if the downsized jobs were picked at random from the two age categories, there is a more than a 5\% chance that the results would be this lopsided. If there is more than a 5\% percent chance of a law-abiding company creating this outcome, courts might be reluctant to infer that a prima facie case for discrimination has been made unless other circumstantial evidence is given. See Mariani & Robertson, \textit{supra} note 142, at 92.

\textsuperscript{147} For a chi-square distribution with $df = 1$, there is a 5\% chance of such an outcome if the
sults of a chi-square test when the age cutoff is moved to the fifty-year mark.

TABLE IV. Comparing employees under fifty to employees fifty and over. The expected value for each category is in parentheses.

<table>
<thead>
<tr>
<th>Employee Age Bands</th>
<th>Number of Employees in Age Band</th>
<th>Number of Employees Fired</th>
<th>Number of Employees Still Employed</th>
</tr>
</thead>
<tbody>
<tr>
<td>under 50</td>
<td>75</td>
<td>10 (15)</td>
<td>65 (60)</td>
</tr>
<tr>
<td>50 and over</td>
<td>25</td>
<td>10 (5)</td>
<td>15 (20)</td>
</tr>
</tbody>
</table>

The chi-statistic here is 6.75. There was a slightly less than one percent chance of this happening if the two age categories were independent of the decisions made to fire people.\(^{148}\)

B. The “Substantially Younger” Requirement and Omitted Data

What would happen if courts, following the O’Connor decision, decide that “substantially younger” means that there must be a five-year gap between the plaintiffs’ subgroup and the group that allegedly benefited from this discrimination?\(^{149}\) The language of the O’Connor decision could mean that plaintiff employees could only be compared statistically to employees who are at least five years younger. In that case the courts might require data to be omitted from the analysis. They might expect the plaintiff to make his prima facie case without including the treatment (whether good or bad) given to other employees who are close to the plaintiff in age.

In Table V, the hypothetical assumes that plaintiffs are all forty-five and older. If they allege that employees under forty-five have been

---

148. There is only 1 degree of freedom here as in Table III. For 1 degree of freedom, chi-square = 3.84 at the 5% level and chi-square = 6.64 at the 1% level. See Curtis, supra note 138, at 221.
149. See e-mail correspondence between the author and Allan G. King, J.D, PhD. (on file with the Hofstra Labor & Employment Law Journal). Dr. King is an attorney with Littler Mendelson, 2001 Ross. Ave., Dallas, TX 75201 (e-mail: AGKing@littler.com). The author would like to thank Dr. King for his help in explaining some of the statistical implications of the O’Connor decision.
given beneficial treatment, the courts might want information on the forty to forty-five year old employees omitted.

TABLE V. Comparison done with a five year gap. Comparing employees thirty-five to forty with those who are forty-five and over. The expected value for each category is in parentheses.

<table>
<thead>
<tr>
<th>Employee Age Bands</th>
<th>Number of Employees in Age Band</th>
<th>Number of Employees Fired</th>
<th>Number of Employees Still Employed</th>
</tr>
</thead>
<tbody>
<tr>
<td>35-40</td>
<td>25</td>
<td>2 (5)</td>
<td>23 (20)</td>
</tr>
<tr>
<td>45-55</td>
<td>50</td>
<td>16 (10)</td>
<td>34 (40)</td>
</tr>
</tbody>
</table>

The chi-square statistic here is 5.34 which is statistically significant. It is less significant than the comparisons done in Tables III and IV, however. This is not surprising, though, because something which reduces the number of observations will usually reduce the chance of getting significant results. In fact, it is a well-known defense strategy to break down the data into as many categories as possible so that significant results become less likely. In the case of Table V, the omitted data which were omitted included the below-average number of layoffs for the forty to forty-five year olds.

For a one-way pattern of numbers...
as was originally shown in Table IA, using the full range of data will probably help the plaintiffs.

Table VI runs an analysis which compares employees in the fifty to fifty-five year range and omits the data for employees forty-five to fifty.

TABLE VI. Comparisons done with a five-year gap. Comparing employees thirty-five to forty-five with employees who are fifty to fifty-five. The expected value for each category is in parentheses.

<table>
<thead>
<tr>
<th>Employee Age Bands</th>
<th>Number of Employees in Age Band</th>
<th>Number of Employees Fired</th>
<th>Number of Employees Still Employed</th>
</tr>
</thead>
<tbody>
<tr>
<td>35-45</td>
<td>50</td>
<td>4 (10)</td>
<td>46 (40)</td>
</tr>
<tr>
<td>50-55</td>
<td>25</td>
<td>10 (5)</td>
<td>15 (20)</td>
</tr>
</tbody>
</table>

The chi-square statistic in Table VI jumps to 8.39 — highly significant but still less significant than when all the data were used in Table IB.

Now look at what happens when all of the chi-square statistics are compared in one chart, listed from greatest significance on top (strongest case for plaintiff) to least significance on bottom (strongest case for defendant). The first column describes the data comparison which created the statistic.

---

153. The odds of this occurring by chance are much less than 1 out of 100. In statistical language, this is written as \( p << 0.01 \).
TABLE VII. A comparison of all the chi-square statistics generated by the comparisons in Tables IB-VI.

<table>
<thead>
<tr>
<th>Data Compared</th>
<th>df</th>
<th>Chi-square Statistic</th>
<th>Significance Level</th>
<th>Original Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>all four</td>
<td>3</td>
<td>11.0</td>
<td>$01 &lt; p &lt; .05$</td>
<td>IB</td>
</tr>
<tr>
<td>age bands</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35-45 vs. 50-55</td>
<td>1</td>
<td>8.39</td>
<td>$p &lt;&lt; .01$</td>
<td>VI</td>
</tr>
<tr>
<td>&lt;45 vs. 45+</td>
<td>1</td>
<td>7.56</td>
<td>$p &lt; .01$</td>
<td>III</td>
</tr>
<tr>
<td>&lt;50 vs. 50+</td>
<td>1</td>
<td>6.75</td>
<td>$p = .01$</td>
<td>IV</td>
</tr>
<tr>
<td>35-40 vs. 45-55</td>
<td>1</td>
<td>5.34</td>
<td>$01 &lt; p &lt; .05$</td>
<td>V</td>
</tr>
<tr>
<td>&lt;40 vs. 40+</td>
<td>1</td>
<td>2.08</td>
<td>(not signif.)</td>
<td>II</td>
</tr>
</tbody>
</table>

The results in Table VII are essentially duplicated even when the statistical test used to analyze the data is different. Table VIII, for example, displays the levels of significance found when the same comparisons done in Tables IB-VI are analyzed with: the Fisher’s Exact Test. Like the chi-square test, the Fisher’s Test sets up contingency tables which are 2X2 and tests to see if the two possible outcomes of the age variable (“younger” or “older”) have a different effect on the job variable (“fired” or “not fired”). The Fisher’s Test is a more powerful test than the chi-square in this situation. As a result, all of the significance levels for the data comparisons are higher (i.e., the $p$ values are lower). The data comparisons are listed in order from “most significant” at the top to “not significant” at the bottom.

154. Read this as “$p$ is much less than .01.”
155. See D. Baldus & J. Cole, STATISTICAL PROOF OF DISCRIMINATION (McGraw Hill, 1980) (§ 9A.11-12); The Fisher’s exact test is used when the sample pool is relatively small. See id. In the case of Table IB, the sample pool is the collection of 100 Downsize Corp. employees who share the same job category. See Carla Walworth & Wendy Dunne DiChristina, Masters of the Numerical Universe: Challenging Statistics in Age Discrimination Suits, EMPLOYMENT REL. L.J. 37, 46 (1995) (recommending the Fisher’s test for small groups).
156. There are two age categories and two job outcomes.
157. The “power” of a statistical test is the probability that it will reject the “null hypothesis” when it is not true. The null hypothesis, in this example, is the theory that older employees have the same chance of getting fired as younger employees. If this is not true, the Fisher’s Test is more likely to point that out. However, there is a tradeoff. More powerful test statistics are also more likely to reject the null hypothesis when they should accept it.
TABLE VIII. A comparison of Fisher's Exact Test statistics generated from the Table IB data. The most significant results are at the top; the least significant are at the bottom.\(^{158}\)

<table>
<thead>
<tr>
<th>Data Compared</th>
<th>Level of Significance</th>
<th>Original Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 45 vs. 45+</td>
<td>( p = 0.0025 )</td>
<td>III</td>
</tr>
<tr>
<td>(most signif.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35-45 vs. 50-55</td>
<td>( p = 0.0057 )</td>
<td>VI</td>
</tr>
<tr>
<td>&lt; 50 vs. 50 +</td>
<td>( p = 0.0063 )</td>
<td>IV</td>
</tr>
<tr>
<td>35-40 vs. 45-55</td>
<td>( p = 0.0182 )</td>
<td>V</td>
</tr>
<tr>
<td>&lt; 40 vs. 40 +</td>
<td>( p = 0.0684 )</td>
<td>II</td>
</tr>
<tr>
<td>(not signif.)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

When the comparisons done with the Fisher's Test are listed in order of significance, the results are almost the same as the chi-square results in Table VII.\(^{159}\) A listing of the different chi-square comparisons from most significant to least significant is as follows: "35-45 vs. 50-55," then "< 45 vs. 45 +," "< 50 vs. 50 +," "35-40 vs. 45-55," and finally "< 40 vs. 40 +." For the Fisher's Exact Test, the most-to-least significant list starts with "< 45 vs. 45 +," then "35-45 vs. 50-55," "< 50 vs. 50 +," "35-40 vs. 45-55," and finally "< 40 vs. 40 +." The chi-square and Fisher's Test produce the same result except that the first two items are reversed.

C. A Legal Interpretation of the Statistical Results\(^{160}\)

Look again at the trend shown in the third column of Table VII.\(^{161}\) These are the results of the chi-square analysis and every comparison that was done produced significant results except for "< 40 vs. 40 +." Still, it is hard to imagine how these statistical results could be translated into legal inferences or conclusions when the results are looked at as a whole. The reason for the difficulty is that the results are not

---

158. Note that the first two items on the list are reversed from the results done with a chi-square statistic, but the remaining data comparisons have retained their significance ranking.

159. The Table IB comparison which included "all four age bands" should not be compared because no analysis was done for all four age bands with the Fisher's Exact Test.

160. Since the trends for both tests were so similar, the rest of this section will concentrate on the chi-square results in Table VII.

161. Since these Tables all involve comparisons with \( df = 1 \), they can be compared to each other. Comparing two chi-square statistics with different degrees of freedom to each other is meaningless, like comparing apples and oranges.
"continuous" — they jump around. To see the discontinuity, focus on the results (in the right-hand column of Table VII) for Tables II, III, and IV. These were the comparisons which did not omit any data; they just involved different cut-off lines separating the employees. Although the original data from Table IIB showed a one-way trend (higher layoffs for higher age groups), the results in Table VII show that the greatest inference of discrimination can be drawn when employees less than forty-five are compared to those over forty-five (Table III). The level of significance there is $p < 0.01$, but falls off when the age cutoff point is either raised or lowered from there. If the cutoff is raised to age fifty the $p$ level rises to 0.01 (Table II). If the age cutoff is lowered to forty, the $p$ level becomes so high it is insignificant (Table II).

This discontinuity has nothing to do with the behavior of the Downsizing Corp. It is a result of a quirk in the mathematics of the statistical test which can not be translated into an easy legal conclusion. For that reason, the plaintiffs might not want to present all of these comparisons to the trier of fact. If the plaintiff did present all of the comparisons, what could a trier of fact conclude from them — that the employer acted in a discriminatory way towards employees up to the age of forty-five, but then treated employees who were even older in a non-discriminatory way? That would not make any sense intuitively or legally. This discontinuity shows why a charge of "data mining" can frequently be leveled against a plaintiff who "subgroups." If the plaintiff is free to select the separation point at any age, then he can make a selection which minimizes the $p$-value and maximizes significance.\(^6\)

The problem can also be seen if the "substantially younger" phrase in O'Connor requires plaintiff to omit data. The plaintiffs' best case in Table VII is when the data for the forty-five to fifty year-olds are omitted ($p << 0.01$). These results are more significant than when the omit-
ted employees are the forty to forty-five year-olds (0.1 < p < 0.05). But Table IA implies that they may have been discriminated against as well. Should the forty to forty-five year olds be excluded from bringing a lawsuit because of the "substantially younger" requirement? Should they be included but treated differently when legal conclusions are drawn?

Not all of the problems are the plaintiffs. Some are faced by the employer. Suppose that the plaintiffs are all between fifty and fifty-five and they have found that their strongest prima facie case (p << 0.01) is when the forty-five to fifty year olds are omitted.163 They have made all of the other comparisons, but see that they are not as convincing. Through the use of attorney-client privilege, they may even keep those results from discovery by the defendant. The company wants to counter the plaintiffs' statistics by showing the actual changes in significance that a different comparison would yield, not just criticizing what the plaintiffs have done. If they have evidence that the gap helps the plaintiff by getting rid of data helpful to the defendant, should they be allowed to show a comparison without any omissions?164 And if the employer does that, how should the plaintiffs be allowed to counter when they try to show "pretext plus"? Courts might want to treat "substantially younger" as a weaker requirement depending on the facts of the case.

VI. SOME POLICY CONCERNS IN APPLYING O'CONNOR

Although O'Connor was decided correctly it is sure to lead to some confusion. Because the Court only ruled what the fourth branch of McDonnell Douglas could not be, but did not say what it must be, lower federal courts will have to work out rules for the "substantially younger" requirement on their own. How these rules evolve will affect the statistical cases which are vital to ADEA "subgrouping" cases. The mathematical example in Section IV shows how the right rules are needed to prevent ADEA plaintiffs from losing more ground than they have already lost under Hicks and Hazen.

163. See the results from Table VI.
164. Or imagine that the results are slightly different from Table VII and that a comparison between the < 45 and 45 + employees give results which are less significant than the results from Table V. Should the employer be allowed to "close the gap" and argue that allowing the gap gives plaintiffs too much opportunity to omit bad data?
To do this, a few simple rules should be followed by the lower courts when applying the "substantially younger" requirement in O'Connor:

1) The meaning of "substantially younger" should vary according to the facts of the case, but as a rule of thumb courts should typically expect a gap of five years between the youngest plaintiff and the oldest beneficiary of the alleged discrimination. This is a large enough gap to keep out frivolous claims without requiring plaintiffs to exclude so much of their data that their cases are harmed by small sample problems.

A five-year gap will also rule out reverse discrimination claims which violate the intent and purpose of the statute. The ADEA was passed to prevent businesses from acting on unsubstantiated stereotypes of older workers and allowing reverse discrimination would completely defeat that purpose. The only time an exception should be made for this is when the plaintiff has presented evidence that the hiring of the older (or same-age) replacement was done as part of a "strategy" to ward off a lawsuit after discriminating on the basis of age against a particular individual.

2) The five-year rule should be relaxed for plaintiffs who have small-sample cases that will be barred by such a rule if those plaintiffs can show that their non-statistical case (involving "stray remarks," discriminatory business practice, etc.) is otherwise strong. The rule should also be relaxed for plaintiffs who can show that statistical anomalies will pop up if they are forced to omit data, but again, this flexibility should be dependent on plaintiffs' showing an otherwise strong case.

3) The employer should be allowed to ignore the "substantially younger" five-year gap when he makes his case for rebuttal. If the defendant counters the prima facie case with a different age comparison (by filling in the missing data, moving the "gap," or running a different "over/under" comparison) the requirement of a five-year difference should be waived for the plaintiffs as well when they try to show "pretext plus."

165. The reasons mentioned here were given by the Seventh Circuit in Hamilton v. Caterpillar Inc., 966 F.2d 1226, 1227 (7th Cir. 1992).

166. Criticism of this sort of "strategic" hire would only make sense if the plaintiff was bringing a "disparate treatment" action where the discrimination was supposed to be intentional. A plaintiff could not allege that an employer's policy was not intentional but had an illegal "disparate impact" on one group of employees at the expense of another group that was the same age or even younger.
Imagine that the example given at the end of Section V is modified slightly: the plaintiffs are all over fifty. They run several different age comparisons and find that their strongest case is when the data for ages forty-three to fifty are omitted. After they make their prima facie case, should the employer be allowed to present results for a comparison of employees under fifty and over fifty? Of course. Limiting the defendant’s methodology to what the plaintiffs originally chose will give plaintiffs too much power. If the defendant’s strongest case is when the “gap” is closed, he should be able to run that test and show the numbers that result. Defendant should not be limited to criticizing the plaintiffs’ methodology. The defendant might even have evidence that “business necessity” made him treat employees fifty and over differently from those “substantially younger.” Then he might want to keep the five-year gap but shift it downward.

And if the employer rebuts the plaintiffs’ methodology with his own, plaintiffs should follow up on the defendant’s rebuttal with any numerical evidence available. The plaintiffs should not be required to show flaws in the defendant’s rebuttal with the same age comparison that they used to make their prima facie case.

These simple rules and their few exceptions will not eliminate the difficulties faced by courts hearing ADEA cases, but they might straighten out some of the confusion that will follow the O’Connor decision. Organizing the theory behind subgroup analysis is important because the American population is aging and faces greater job insecurity from cost-cutting employers. There are going to be many such employers who take the decision-making “short cut” of simply firing their oldest workers to save money without considering the actual benefits.

167. See Carla Walworth & Wendy DiChristina, *Master of the Numerical Universe: Challenging Statistics in Age Discrimination Suits*, 21 EMPLOYMENT REL. L.J. 37 (1996) for a discussion of the defendants’ tactics in light of possible statistical quirks. “[P]laintiff may proffer statistical evidence supporting an inference that an employer discriminated against workers over 50 years of age. If, however, the employer can establish that workers 55 years of age and older are not discriminated against, then the plaintiff’s statistical evidence could be merely an arbitrary anomaly . . . .” *Id.* at 44.
As Judge Pierce pointed out in his concurring decision in Lowe, these future plaintiffs will be replaced by people within the protected category and it will help to have rules for "substantially younger" settled when they plan their prima facie cases.

Kurt Schaub *

---

* I would like to acknowledge the love and patience of my wife, Karen Elise Schaub. Without her support, this Note could not have been written.