2010

Gross Disunity

Martin J. Katz
University of Denver, mkatz@law.du.edu

Gross Disunity

Publication Statement

Originally published as Martin J. Katz, *Gross Disunity*, 114 PENN. ST. L. REV. 857 (2010).

Copyright is held by the author. User is responsible for all copyright compliance.

PENN STATE LAW REVIEW

Gross Disunity

Martin J. Katz*

Table of Contents

INTRODUCTION

The Supreme Court has done a turn-about on the value of uniformity in employment discrimination law. For many years, the Court embraced the idea that different employment discrimination statutes that use identical language should be understood to impose identical requirements.[1] So, for example, a plaintiff claiming age

* Dean and Associate Professor of Law, University of Denver College of Law; Yale Law School, J.D. 1991; Harvard College, A.B. 1987. Thanks to the Colorado Employment Law Faculty (Rachel Arnow-Richman, Roberto Corrada, Scott Moss, Melissa Hart, Raja Raghunath, Helen Norton, and Nantiya Ruan) for their comments on drafts. Thanks to Blair Kanis for her research assistance. Any errors are my own.

1. When I talk about the requirements for proving discrimination, I refer to definitions of causation—that is, what type of causation a party must prove to prevail on a claim or defense. Several commentators in this area refer to "proof structures." *See, e.g.*, William R. Corbett, *Fixing Employment Discrimination Law*, 62 SMU L. REV. 81, 90 (2009); Charles A. Sullivan, *Disparate Impact: Looking Past the* Desert Palace

discrimination under the Age Discrimination in Employment Act (ADEA)[2] would face the same requirements as a plaintiff claiming race or sex discrimination under Title VII of the Civil Rights Act of 1964 (Title VII).[3] More recently, the Court has moved away from this ideal of uniformity. And last summer, in *Gross v. FBL Financial Services*,[4] the Court completely rejected that ideal.

This Article will argue that the Court's rejection of unification in this field is normatively problematic for four reasons:[5] First, in most instances, uniformity is desirable, both as a matter of efficient legal administration and as an assumption about Congressional intent. *Gross* eschews these benefits without an explanation of why age discrimination should be treated differently than race or sex discrimination, or why Congress might have wanted to treat these forms of discrimination differently.

Second, the timing of the Court's rejection of uniformity looks bad. The Court embraced uniformity during a time when doing so had the effect of expanding the application of its own definition of

Mirage, 47 WM. & MARY L. REV. 911, 912 (2005); Michael J. Zimmer, *The New Discrimination Law:* Price Waterhouse *is Dead, Whither* McDonnell Douglas?, 53 EMORY L.J. 1887, 1891 (2004); Deborah C. Malamud, *The Last Minuet: Disparate Treatment after* Hicks, 93 MICH. L. REV. 2229, 2229 (1995). By "proof structures," these authors refer to paths for proving discrimination set out by different Supreme Court opinions and statutes, such as the *McDonnell Douglas* proof structure or the *Price Waterhouse* proof structure. The problem with this approach is that some of these opinions and statutes set out specific causal standards, while others set out procedures that may be used to prove specific causal standards. *See* Martin J. Katz, *Reclaiming* McDonnell Douglas, 83 NOTRE DAME L. REV. 109, 143 (2007). To avoid this ambiguity, I will refer to specific definitions of causation, rather than proof structures.

2. 29 U.S.C. §§ 621-34 (2000).

3. 42 U.S.C. §§ 2000e-2000e-17 (2000).

4. Gross v. FBL Fin. Servs., 129 S. Ct. 2343 (2009).

5. Other authors who have criticized *Gross* include William R. Corbett, *Babbling about Employment Discrimination Law: Does the Builder Understand the Blueprint for the Great Tower*, 26 (forthcoming in U. PENN. J. BUS. & EMPL. L. 2010) (manuscript on file with Author); David G. Savage, *Age bias much harder to prove: The Supreme Court shifts the burden of proof to the worker making the claim. Businesses cheer*, L.A. TIMES, June 19, 2009, at 1, *available at* http://bulletin.aarp.org/yourworld/law/articles/supreme_court_makes_agebias_suits_harder_to_win.html; Editorial, *Age Discrimination*, N.Y. TIMES, July 7, 2009, at A22 (calling for Congress to reverse *Gross* as it did *Ledbetter*); Kevin P. McGowan, *EEOC Provides Guidance on Waivers, Hears Testimony on Age Bias Developments*, 134 Daily Lab. Rep. (BNA) A-14 (July 16, 2009); Susan J. McGolrick, *Justices 5-4 Adopt But-For Causation, Reject Burden Shifting for ADEA Claims*, 116 Daily Lab. Rep. (BNA) AA-1 (June 19, 2009). Senate Judiciary Committee Chairman Senator Patrick Leahy stated as follows: "By disregarding congressional intent and the time-honored understanding of the statute, a five member majority of the Court has today stripped our most senior American employees of important protections." *Id.* Senator Leahy further likened the *Gross* decision to the Court's "wrong-headed" ruling in *Ledbetter v. Goodyear Tire & Rubber Co*, 550 U.S. 618 (2007), which Congress overturned in the Lilly Ledbetter Fair Pay Act of 2009. *Id.*

discrimination. But when Congress supplanted the Court's definition with one of its own in 1991, the Court seemed to sour on unification. This timing may suggest recalcitrance.

Third, the Court went further than required under its reasoning, unnecessarily rejecting a burden-shifting mechanism that is important for plaintiffs. The Court's argument against unification may have supported a rejection of Congress's 1991 Title VII definition of discrimination in ADEA cases. However, the Court went a step further, also rejecting the application of its own pre-1991 Title VII definition of discrimination—the definition set out in *Price Waterhouse v. Hopkins*.[6] Doing so had the effect of eradicating burden-shifting from the ADEA.[7]

Fourth, *Gross* adopted a normatively problematic definition of discrimination: It required plaintiffs in ADEA cases to prove but-for causation without the aid of a burden-shifting mechanism. Among the causal standards available in modern employment discrimination law, this standard is the worst. Moreover, *Gross*'s reasoning suggests that the Court is likely to apply this normatively problematic standard to all employment discrimination statutes other than the part of Title VII that was amended by the Civil Rights Act of 1991.[8]

This Article will proceed as follows: Part I will explain *Gross* in terms of causation and unification. Part II will argue that *Gross* rejected the doctrine of uniformity, a well-established and useful canon of statutory construction, without explanation. Part III will show how the courts' post-1991 rejection of uniformity, culminating in *Gross*, might be seen as a form of judicial recalcitrance. However, that Part will suggest that the Court's rejection of uniformity in *Gross* is better understood as a

6. Price Waterhouse v. Hopkins, 490 U.S. 228 (1989).

7. There is another type of burden-shifting. The framework in *McDonnell Douglas Corp. v. Green* shifts a burden to the defendant. *See* 411 U.S. 792, 800 (1973). However, *McDonnell Douglas* only shifts the burden of production (of articulating a non-discriminatory reason for the decision in question)—not the burden of persuasion. *See id.* Accordingly, in this article, I will not refer to *McDonnell Douglas* as a burden-shifting framework. I will reserve that label for frameworks that shift the burden of persuasion, such as *Price Waterhouse*.

8. Some courts have already done this. *See, e.g.*, Fairly v. Andrews, 578 F.3d 518 (7th Cir. 2009) (applying *Gross* to First Amendment free speech claim); Levi v. Wilts, No. 08-3042, 2009 WL 2905927 (C.D.Ill. Sept. 4, 2009) (applying *Gross* to a constitutional retaliation claim); Williams v. District of Columbia, 646 F. Supp. 2d 103 (D.D.C. 2009) (applying *Gross* to a claim brought under the Juror Act). *See also* Postings of Paul M. Secunda & Steve Kaminshine, *Zimmer on* Gross *ADEA Case and Employer Strategy*, to Workplace Prof Blog, http://lawprofessors.typepad.com/laborprof_blog/2009/11/zimmer-on-gross-adea-case-and-employer-strategy.html (Nov. 4, 2009) (quoting Posting of Michael Zimmer, *The Employer's Strategy in* Gross v. FBL Financials, to Concurring Opinions, http://www.concurringopinions.com/?s=gross+adea+and+employer (Nov. 4, 2000, 10:43 EST)) (speculating that courts will likely apply *Gross* to cases under other disparate treatment statutes).

rejection of burden-shifting in disparate treatment doctrine. Finally, Part IV will argue that burden-shifting is normatively desirable in disparate treatment doctrine, and that *Gross* adopted the worst of the causal standards available to it. The Article concludes with a call for decisive legislative action.[9]

I. *GROSS*, CAUSATION, AND UNIFICATION

Almost all disparate treatment statutes include an element of causation.[10] They do not prohibit adverse employment actions, such as firing, in all instances. Rather, they prohibit adverse employment actions only where those actions occur "because of" a protected characteristic, such as race, sex, or age.[11] In other words, these statutes all require causation.

Virtually all of these disparate treatment statutes use the same phrase—"because of"—to describe their causation requirement.[12] Yet, until 1991, none of those statutes actually defined this phrase. And there are several possible meanings for the phrase.[13]

A. The Ambiguity in "Because of"

To understand the ambiguity in the phrase "because of," it is helpful to imagine an adverse employment decision, such as employer's decision to fire an employee. And imagine that a protected factor, such as the employee's race, played a role in that decision. The question is exactly what role the protected factor played in the decision. There are four possibilities; that is, there are four types of causation:

9. As of the time this Article went to print, a bill entitled the Protecting Older Workers Against Discrimination Act has been introduced in Congress. *See* H.R. 3721, 111th Cong. (2009); S. 1756, 111th Cong. (2009). This bill will be discussed more fully below. *See infra* notes 137-1378 and accompanying text.

10. There are two basic types of anti-discrimination law: disparate treatment law (which involves so-called intentional discrimination) and disparate impact law (which involves statistical disparities that may or may not be caused by intentional discrimination). This Article focuses on disparate treatment law.

11. *See, e.g.*, Title VII of the Civil Rights Act of 1964, as amended, 42 U.S.C. § 2000e (prohibiting adverse employment actions where they occur because of race, color, national origin, religion, or sex); Age Discrimination in Employment Act ("ADEA"), 29 U.S.C. §§ 623-633a (2000) (prohibiting adverse employment actions where they occur because of age); Americans with Disabilities Act ("ADA"), 42 U.S.C. §§ 12101-12117 (2000) (prohibiting adverse employment actions where they occur because of disability).

12. See statutes cited in *supra* note 11.

13. The majority in *Gross* says that there is only one ordinary meaning for this phrase: but-for causation. *See* Gross v. FBL Fin. Servs., 129 S. Ct. 2343, 2350 (2009). However, as the text immediately below shows, the majority was wrong.

- Necessity (often called "but-for" causation). A factor (consideration of race) is necessary to an outcome (the decision to fire) where the outcome would not have occurred absent—or but for—that factor.

- Sufficiency. A factor (consideration of race) is sufficient where, given all of the other factors then present, adding that factor will inevitably trigger the outcome (the decision to fire).

- Minimal Causation (often called "motivating factor" causation). A factor (consideration of race) is minimally causal—a motivating factor—where that factor has a tendency to affect the outcome (the decision to fire). A factor can be minimally causal without being either necessary or sufficient. So minimal/motivating factor causation is less restrictive than either necessity/but-for causation or sufficiency.

- Sole Causation. A factor (consideration race) is the sole cause of an outcome (the decision to fire) where there are no other minimally causal factors present. Sole causation is the most restrictive causal concept. A factor that is a sole factor will also satisfy any other type of causation requirement.[14]

Because the first two types of causation (necessity and sufficiency) can be combined in two different ways to form a causation requirement ("necessity *and* sufficiency" or "necessity *or* sufficiency") there are actually six potential causation requirements—six potential meanings for the phrase "because of." But even this is an oversimplification. This is because, in deriving a causation requirement, it is possible to require one type of causation for one purpose and another type of causation for another purpose. For example, the phrase "because of" might require minimal/motivating factor causation for liability, but require necessity/but-for causation for full damages.[15] So the phrase "because of" is an ambiguous one.

14. A factor which is the sole cause will also be minimally causal, necessary, and sufficient. However, a factor can be minimally causal, necessary, and sufficient even where there are other minimally causal factors present. So, sole causation is distinct from these other three concepts. In an earlier work, I equated sole causation with a combination of necessity and sufficiency. *See* Martin J. Katz, *The Fundamental Incoherence of Title VII: Making Sense of Causation in Disparate Treatment Law*, 94 GEO. L.J. 489 (2006). The preceding logic shows that this was mistaken.

15. *See, e.g.*, Civil Rights Act of 1991, Pub. L. No. 102-166, § 107(a), 105 Stat. 1071, 1075 (codified at 42 U.S.C. § 2000e-2(m) (2000)) (providing that plaintiff must show that a protected characteristic was a "motivating factor" in the adverse decision); *Id.*

The Court and most commentators have tended to focus on the first and third definitions of causation: necessity/but-for causation, and minimal/motivating factor causation.[16] But even with only these two potential definitions, the phrase "because of" is ambiguous. "Because of" could mean either of these types of causation, or some combination of them (such as minimal/motivating factor causation for liability, and necessity/but-for causation for full damages).

B. The Meaning of "Because of" in Title VII

Prior to 1989, the Court did not definitively address the meaning of the phrase "because of" in Title VII or any of the major disparate treatment statutes. Arguably, the Court had no need to do so, for it had yet to encounter a true multi-factor case. The pre-1989 cases it saw were either-or cases. In those cases, one party had argued that the adverse employment decision in question was caused by one factor (such as race) and the other party had argued that the decision was caused by another factor (a non-discriminatory factor, such as insubordination).[17] Thus, the Court had to decide only which one of these two factors caused the decision. If only one factor causes a decision, then it will be a sole cause. And when something is a sole cause, it will satisfy every other conceivable causation requirement. Accordingly, there was no need for the Court to decide the meaning of "because of."

In 1989, in *Price Waterhouse v. Hopkins*,[18] the Court was faced, for the first time, with a case that required it to define "because of"—a multi-factor case. In that case, the defendant claimed that, even if it had used a discriminatory factor (sex) in its decision-making, it would have reached the same decision (non-promotion) based on another, independently sufficient factor ("abrasiveness").[19] This type of case

at § 107(b)(3), 105 Stat. at 1075-76 (codified at 42 U.S.C. § 2000e-5(g)(2)(B) (2000)) (providing that once plaintiff has done so, defendant may demonstrate that it would have taken the "same action" absent consideration of the protected characteristic).

16. See statutes cited *supra* note 15.

17. *See, e.g.*, Tex. Dep't of Cmty. Affairs v. Burdine, 450 U.S. 248 (1981); McDonnell Douglas Corp. v. Green, 411 U.S. 792, 800 (1973).

18. Price Waterhouse v. Hopkins, 490 U.S. 228 (1989).

19. *See id.* at 234 (plurality opinion). There was some question about whether "abrasiveness" was actually independent from sex—that is, whether an employee at Price Waterhouse could or would be perceived as "abrasive" independently of sex; or whether only women tended to be given this label. On remand, the lower court found that sex and abrasiveness were not, in fact, independent at Price Waterhouse. *See* Hopkins v. Price Waterhouse, 737 F. Supp. 1202, 1207 (D.D.C.), *aff'd*, 920 F.2d 967 (D.C. Cir. 1990). (For this reason, I use quotations around the word.) However, for purposes of its decision, the Supreme Court assumed that "abrasiveness" might be independent from sex, thereby creating the possibility of multiple, independent causal factors—and thus the need to define "because of."

squarely presented the question of which type of causation was required by the phrase "because of." In such a case, sex discrimination could be a motivating factor in, yet not be a but-for cause of, the adverse decision.

Price Waterhouse held that "because of" in Title VII has two meanings for two different purposes. The Court held that a plaintiff who showed motivating factor causation could shift the burden of persuasion to the defendant.[20] But the Court held that the defendant can then avoid liability by showing that it would have reached the same decision irrespective of its use of the protected motivating factor (sex).[21] This same decision formulation is a but-for test; it requires the defendant to prove a lack of but-for causation.[22] By doing so, the defendant avoids liability. So *Price Waterhouse* requires but-for causation for liability, but only motivating factor causation for burden-shifting.

There was some question as to whether *Price Waterhouse* contained an additional requirement for burden-shifting—above and beyond proving motivating factor causation. The plurality of four Justices did not require anything else to shift the burden.[23] Nor did Justice White's concurrence.[24] So these five Justices can be seen as adopting a simple, motivating factor standard for burden-shifting.[25] However, Justice O'Connor, whose concurrence in *Price Waterhouse* is generally seen as controlling,[26] added one additional requirement for burden-shifting: The plaintiff must prove motivating factor causation by "direct evidence."[27] If the plaintiff does not have "direct evidence," she must prove but-for causation with no burden-shifting.[28]

20. *See Price Waterhouse*, 490 U.S. at 248 (plurality opinion). Portions of the opinion used the phrase "motivating factor" and "substantial factor" interchangeably. *See* Katz, *supra* note 14. In an earlier article, I demonstrated that these two phrases should be seen as synonymous. *See id.*

21. *See Price Waterhouse*, 490 U.S. at 248 (plurality opinion).

22. *See* Katz, *supra* note 14. The plurality in *Price Waterhouse* disclaimed the but-for standard. *See* 490 U.S. at 240 (plurality opinion) ("To construe the words 'because of' as colloquial shorthand for 'but-for causation,' . . . is to misunderstand them."). However, as demonstrated in the text above, the plurality's same decision defense adopts a but-for standard for liability. *See id.* at 281 (Kennedy, J., dissenting) ("The theory of Title VII liability the plurality adopts, however, essentially incorporates the but-for standard.").

23. *See Price Waterhouse*, 490 U.S. at 244 (plurality opinion).

24. *See id.* at 259 (White, J., concurring in the judgment).

25. *See Gross*, 129 S. Ct. at 2357 (Stevens, J., dissenting) (suggesting that Justice White's concurrence might be controlling opinion in *Price Waterhouse*); *see also* Martin J. Katz, *Unifying Disparate Treatment (Really)*, 59 HASTINGS L.J. 643, 661 (2008) (same).

26. *See* Zimmer, *supra* note 1, at 1910; *see also* Griffith v. City of Des Moines, 387 F.3d 733, 743 (8th Cir. 2004) (Magnuson, J., concurring specially) (noting that after *Price Waterhouse*, courts follow Justice O'Connor's concurrence).

27. *See Price Waterhouse*, 490 U.S., at 276 (O'Connor, J., concurring in the judgment). I put "direct evidence" in quotes because its meaning is unclear. *See* Costa v.

So after *Price Waterhouse*, in Title VII cases, "because of" meant: (1) motivating factor for burden-shifting in cases with "direct evidence," and (2) but-for for liability in all cases.

Following *Price Waterhouse*, Congress adopted a different—and less restrictive—definition of "because of" for Title VII. In 1991, Congress rejected but-for causation as a liability standard. It amended Title VII to make clear that "because of" means (1) motivating factor causation for both liability and burden-shifting, and (2) but-for causation for an award of full damages.[29] While there was initially some question about whether a Title VII plaintiff needed to prove motivating factor causation by "direct evidence," the Court in *Desert Palace v. Costa* unanimously concluded that there was no such requirement.[30]

C. To Unify or Not to Unify: "Because of" in the ADEA

So the Court, and later Congress, defined "because of" in Title VII, which deals with discrimination on the basis of race, color, sex, national origin, or religion.[31] But the plaintiff in *Gross* did not fit within one of these categories. Rather, Jack Gross claimed that he was demoted based on his age; so his claim was brought under the Age Discrimination in Employment Act (ADEA).[32] Accordingly, the Court in *Gross* had to figure out the meaning of "because of" under the ADEA.

Desert Palace, 299 F.3d 838, 852-53 (9th Cir. 2002), *aff'd*, 539 U.S. 90 (2003) (explaining various definitions of "direct evidence").

28. More precisely, Justice O'Connor said that if the plaintiff does not have "direct evidence," then she must prove discrimination using the three-step procedure set out in *McDonnell Douglas v. Green*, 411 U.S. 792, 800 (1973). *See Price Waterhouse*, 490 U.S., at 276 (O'Connor, J., concurring). However, most courts and commentators understand *McDonnell Douglas* as a simple but-for standard; that is, as a requirement that the plaintiff bear the full burden of proving but-for causation. *See* Katz, *supra* note 1. In that article, I argue that it is a mistake to equate *McDonnell Douglas* with but-for causation. *See id.* However, that issue is not important to this Article. Accordingly, in this Article, I will simply accept that many courts and commentators do equate *McDonnell Douglas* and but-for causation.

29. 42 U.S.C. § 2000e-2(m) (2000) (providing that plaintiff must show that a protected characteristic was a "motivating factor" in the adverse decision); 42 U.S.C. § 2000e-5(g)(2)(B) (providing that once plaintiff has done so, defendant may demonstrate that it would have taken the "same action" absent consideration of the protected characteristic).

30. *See Desert Palace*, 539 U.S. at 101. This may be a slight oversimplification. In a footnote, the Court stated that it was only addressing "mixed motive" cases. *Id.* at 94 n.1. This reservation might be read to suggest that the "direct evidence" test may continue to apply in 1991 Act cases that are not "mixed motive" cases—that in "single motive" cases without "direct evidence," plaintiffs must use the *McDonnell Douglas* framework. However, for reasons I explain elsewhere, this more complicated approach makes no sense. *See* Katz, *supra* note 1, at 135 n.115.

31. *See* 42 U.S.C. §2000e-2(a).

32. *See* Gross v. FBL Fin. Servs., 129 S. Ct. 2343, 2347 (2009).

This issue, in turn, presented a question of unification: Should the Court assume that "because of" in the ADEA has the same meaning as it does in Title VII?

The question of unification in *Gross* could have been complicated by the passage of the 1991 Act. If the phrase "because of" in the ADEA has the same meaning as it does under Title VII, it raises the question: Does it have the same meaning as it does under pre-1991 Title VII (the *Price Waterhouse* definition) or post-1991 Title VII (the 1991 Act definition)? The latter view—that the 1991 Act definition applies in non-Title VII statutes—is a *total unification* position. If we adopt this position, a single definition (the 1991 Act definition) would apply in all disparate treatment cases.[33] On the other hand, the former view—that the pre-1991 Title VII definition (*Price Waterhouse*) applies in non-Title VII statutes—is a *partial unification* position. If we adopt this position, there would be two regimes: Title VII cases would use the 1991 Act definition,[34] while all other disparate treatment statutes would use the *Price Waterhouse* definition. That is, there would be unification among all non-Title VII cases (which would all use the *Price Waterhouse* definition), but not between Title VII and non-Title VII cases.

However, in *Gross* none of the parties argued for total unification.[35] So the Court was not required to choose between total unification and partial unification. Rather, it was only required to choose between partial unification (application of *Price Waterhouse* to all non-Title VII cases, including ADEA cases) and non-unification (application of some standard other than *Price Waterhouse* to ADEA cases).

Notably, both of the parties in *Gross*, and both lower courts, seemed to assume partial unification. They all assumed that to determine the proper definition of "because of" under the ADEA they needed to look to the same phrase in pre-1991 Title VII—which had been defined by *Price Waterhouse*.

33. I made an argument for total unification in *Unifying Disparate Treatment (Really)*, *supra* note 25.

34. This may be a slight oversimplification. In the partial unification view that has been adopted by the courts, the 1991 Act definition does not apply to all of Title VII. Rather, it applies only to Section 703(a) of Title VII—the part of Title VII that was amended by the relevant portion of the 1991 Act. Other parts of Title VII, such as its anti-retaliation provision, are often understood as using the pre-1991 definition of "because of." *See* Katz, *supra* note 25. Because the distinction is not directly relevant to this Article, I will often simplify by distinguishing Title VII from other statutes, rather than Section 703(a) of Title VII from other statutes.

35. This was likely because most of the lower courts had rejected the total unification option. *See infra* note 778 and accompanying text. *See also* Lawhead v. Ceridian Corp., 463 F. Supp. 2d 856 (N.D. Ill. 2006) (surveying case law on applicability of Title VII standard in non-Title VII cases).

The only dispute between the parties in *Gross* was whether the *Price Waterhouse* definition should apply fully or only partially. Specifically, the fight was over whether the "direct evidence" requirement from *Price Waterhouse* applied in ADEA cases. The defendant argued for the application of the full *Price Waterhouse* definition, including its "direct evidence" requirement (motivating factor causation for burden-shifting only in cases with "direct evidence," and but-for causation for liability).[36] The plaintiff argued for the application of the *Price Waterhouse* definition without the "direct evidence" requirement (motivating factor causation for burden-shifting irrespective of "direct evidence," and but-for causation for liability).[37] The District Court sided with the plaintiff, and the Court of Appeals sided with the defendant.[38] But all assumed that *Price Waterhouse* applied to the ADEA—that the definition of "because of" in the ADEA was tied to the definition of that same phrase in pre-1991 Title VII.

The Supreme Court initially appeared to make the same assumption. The Court granted certiorari to determine "whether a plaintiff must present direct evidence of age discrimination in order to obtain a mixed-motives [burden-shifting] jury instruction in a suit brought under the [ADEA]."[39] In other words, like the parties and the courts below, the Supreme Court appeared to accept that some version of *Price Waterhouse* applied in ADEA cases.

Yet the majority in *Gross* went on to ignore the question posed in its grant of certiorari and to reject any assumption of unification. Rather, the Court, in a 5-4 decision, concluded that "because of" in the ADEA means something completely different—and more restrictive—than it does under any of the two versions of *Price Waterhouse* advanced by the parties. The *Gross* Court concluded that, under the ADEA, "because of"

36. *See Gross*, 129 S. Ct. at 2348.

37. *See id*. The plaintiff needed *Price Waterhouse* for his burden-shifting argument. So he argued that, while *Price Waterhouse* applies in ADEA cases, its "direct evidence" requirement does not apply to the ADEA because the ADEA contains no mention of "direct evidence." *See id*.; *see also* Desert Palace v. Costa, 539 U.S. 90 (2003) (holding that there is no "direct evidence" requirement in post-1991 Title VII because that statute does not mention "direct evidence"); Rachid v. Jack in the Box, 376 F.3d 305 (5th Cir. 2004) (applying same argument to ADEA). Of course, the plaintiff in *Gross* sought to rely on pre-1991 Title VII, which had been interpreted by *Price Waterhouse* as including a "direct evidence" requirement. But my point is not that the plaintiff's argument was perfectly consistent. My point is that the plaintiff, like the defendant, relied on *Price Waterhouse*, a Title VII case, to define the phrase "because of" in the ADEA.

38. *See Gross*, 129 S. Ct., at 2348.

39. *See id.* at 2346.

means but-for causation with no burden-shifting.[40] In the following parts, I will show why this choice was problematic.[41]

II. *GROSS*'S REJECTION OF UNIFORMITY

A well established canon of construction suggests that, in cases like *Gross*, where Congress has used the same phrase in two similar statutes, courts should use the same definition for the phrase in both statutes.[42]

The unification canon has some practical benefits. Courts and litigants must deal with a multitude of disparate treatment statutes. Title VII deals with race, color, national origin, religion and sex. The ADEA deals with age. Still other statutes deal with disability, family leave status, veteran status, and various other protected criteria.[43] Using a single standard in all of those statutes would simplify this terrain; it would allow parties and litigants to resolve issues under any particular statute by reference to a common body of law applicable to several statutes.[44] This would be particularly helpful in cases involving claims under multiple statutes. For example, in a case where the plaintiff claims discrimination on the basis of sex and age, a unified standard would

40. *See id.* at 2351.

41. Several commentators have expressed surprise at *Gross*'s rejection of *Price Waterhouse*. *See* Corbett, *supra* note 5, at 14 (forthcoming in U. PENN. J. BUS. & EMPL. L. 2010) (manuscript on file with Author); Melissa Hart, *Procedural Extremism: The Supreme Court's 2008-2009 Labor and Employment Decisions*, 13 EMPLOYEE RTS. & EMP. POL'Y J. 253 (forthcoming 2009) (manuscript on file with author). Apparently, these authors also expected the Court to adopt a partial unification position.

42. *See* WILLIAM N. ESKRIDGE, JR. ET AL., LEGISLATION AND STATUTORY INTERPRETATION (2d ed. 2006); 2B SUTHERLAND ON STATUTORY CONSTRUCTION, §§ 51:1-8 (7th ed.); William N. Eskridge, Jr. et al., *Forward: Law as Equilibrium*, 108 HARV. L. REV. 26 (1994) (and cases cited therein). *See also* Northcross v. Memphis Bd. of Educ., 412 U.S. 427, 428 (1973) (stating that similarity in language in two statutes "is, of course, a strong indication that the two statutes should be interpreted pari passu."); Jamie Darin Prenkert, *Bizarro Statutory Construction*, 28 BERKELEY J. EMP. & LAB. L. 217, 234-35 (2007) ("When the legislature borrows language from one statute to draft a subsequent statute, courts generally agree that the statutes should be construed consistently.").

43. *See, e.g.*, Americans with Disabilities Act (ADA), 42 U.S.C. § 12101 et seq. (disability); Family and Medical Leave Act (FMLA), 29 U.S.C. § 2601 et seq. (family leave status); 29 U.S.C. § 4301 et seq. (veteran status).

44. *See* Corbett, *supra* note 5, at 8 (forthcoming in U. PENN. J. BUS. & EMPL. L. 2010) (manuscript on file with author) ("A high degree of symmetry among the various laws and covered characteristics may also be desirable, as this may improve simplicity and certainty. . . ."). Professor Corbett goes on to note that "complete uniformity" may not be appropriate "because discrimination based on the various protected characteristics is not a monolithic phenomenon, and the goals of and rationales for the laws differ somewhat." *See id.* (footnotes omitted). I address the potential differences between the ADEA and Title VII below. *See infra* notes 501-523 and accompanying text.

permit the judge and jury to apply a single standard to all of the claims—as opposed to having to apply different standards to different claims.[45]

But the primary strength of the unification canon is its grounding in powerful assumptions about Congressional intent. Where Congress uses a phrase in an earlier statute and then uses that exact same phrase again in a later statute on the same topic, it is generally reasonable to presume that Congress intended the phrase to have the same meaning in the later statute as it did in the earlier statute.[46]

In Title VII, in 1964, Congress used the words "because of" to preclude employment decision-making based on race, color, sex, national origin, or religion.[47] Three years later, in 1967, Congress used those same words, "because of," to preclude employment decision-making based on age in the ADEA.[48] All other things being equal, it would make sense to assume that Congress meant the phrase to mean the same thing in the ADEA as it meant in Title VII.[49]

Of course, all other things may not be equal. There may be times when the presumption of uniformity does not make sense. For example, if Congress uses two different phrases in two similar statutes, it makes sense to assume that Congress intended those phases to have different meanings.[50] Similarly, when Congress uses different remedial or procedural provisions in two similar statutes, it makes sense to assume that Congress intended different interpretations of those remedial or procedural provisions.[51] However, these exceptions support the rule: Just as it makes sense to assume that Congress intended different provisions to be interpreted differently, it makes sense to assume that Congress intended identical provisions—such as the phrase "because of" in Title VII and the ADEA—to be interpreted identically.

45. *See Gross*, 129 S. Ct. at 2357 (Stevens, J., dissenting) (majority's standard "will further complicate every case in which a plaintiff raises both ADEA and Title VII claims.").

46. *See* Prenkert, *supra* note 42, at 234-35.

47. *See* 42 U.S.C. § 2000e-1(a).

48. *See* 29 U.S.C. § 623(a).

49. Prior to *Gross*, most courts to address this issue adopted this presumption of unification in ADEA cases, as well as in other disparate treatment cases. *See infra* notes 75-77 and 80. *See also Gross*, 129 S. Ct. at 2355 (Stevens, J., dissenting) ("The relevant language in the two statutes is identical, and we have long recognized that our interpretations of Title VII apply 'with equal force in the context of age discrimination. . . .'").

50. *See, e.g.*, Meachum v. Knolls Atomic Power Lab., 128 S. Ct. 2395 (2008) (holding that different language describing defenses in Title VII and ADEA suggests that Congress intended different interpretations of those defenses).

51. *See, e.g.*, Lorillard v. Pons, 434 U.S. 575, 584 (1978) (finding significant differences in remedial and procedural provisions of Title VII and ADEA).

It also might make sense to abandon the presumption of uniformity if there were reason to believe that Congress thought that the problem it was trying to address in one statute differed in some important way from the problem it was trying to address in the other. For example, suppose that Congress thought race or sex discrimination were more prevalent or more pernicious than age discrimination.[52] If that were the case, Congress might choose to make it easier to prove race or sex discrimination than age discrimination by adopting a less restrictive definition of "because of" in Title VII and a more restrictive definition in the ADEA. However, if the Court were to conclude that Congress wanted to privilege Title VII claims over ADEA claims in this way, we would generally expect the Court to make this argument expressly—and to provide evidence of such Congressional thinking. Yet, the *Gross* Court offers no such reasoning and no such evidence to support its rejection of the presumption of uniformity.

52. *See, e.g.*, Corbett, *supra* note 1, at 90 ("[T]he Supreme Court consistently has said that there are differences between age discrimination on the one hand, and race and sex discrimination on the other."). Such differences might include, for example, assumptions about the severity of discrimination faced by each group. *See, e.g.*, Hazen Paper Co. v, Biggins, 507 U.S. 604, 610 (1993) (suggesting that discrimination against older workers tends to involve stereotyping, rather than animus—but nevertheless assuming uniformity between the ADEA and Title VII). Or such differences might involve assumptions about the relative prevalence of each type of discrimination. *See, e.g.*, Michael Selmi, *Why are Employment Discrimination Cases So Hard to Win?*, 61 LA. L. REV. 555, 564-65 (2001) (suggesting that courts might be skeptical about the prevalence of discrimination against workers at the lower end of the protected age range); Deborah A. Calloway, *St. Mary's Honor Center v. Hicks: Questioning the Basic Assumption*, 26 CONN. L. REV. 997, 1036 (1994) (suggesting that courts are increasingly skeptical about the prevalence of race discrimination). Alternatively, the Court might take the position that some rights are more "deserving" of protection than other rights. *See, e.g.*, Postings of Michael Zimmer & Steve Kaminshine, Workplace Prof Blog (Nov. 4, 2009) (http://lawprofessors.typepad.com/laborprof_blog/2009/11/zimmer-on-gross-adea-case-and-employer-strategy.html). As noted in the text below, the *Gross* Court rejected the presumption of uniformity without a discussion about any of the potential differences between age discrimination and race or sex discrimination.

The legislative history of the ADEA suggests that Congress considered, but eventually rejected, the possibility that age discrimination was less problematic than race or sex discrimination. When it passed Title VII, Congress considered adding age to the list of protected criteria in that statute. *See* 110 Cong. Rec. 2596-2599 (1964) (amendment offered by Rep. Dowdy, voted down 123 to 94); *id.* at 9911-9913, 13490-13492 (amendment offered by Sen. Smathers, voted down 63 to 28). Instead of adding age to Title VII, Congress requested the Secretary of Labor to study the problem of age discrimination and to make a recommendation to Congress. § 715, 78 Stat. 265. The Secretary reported that age discrimination was indeed a problem. REPORT OF THE SECRETARY OF LABOR, THE OLDER AMERICAN WORKER: AGE DISCRIMINATION IN EMPLOYMENT 5 (June 1965), *reprinted in* U.S. EQUAL EMPLOYMENT OPPORTUNITY COMMISSION, LEGISLATIVE HISTORY OF THE AGE DISCRIMINATION IN EMPLOYMENT ACT (1981) DOC. NO. 5. Following this Report, Congress passed the ADEA, using the same "because of" language as it had used in Title VII.

The only argument that the *Gross* Court offers for rejecting the presumption of uniformity is what I call the *limited amendment* argument: When Congress amended Title VII in 1991, setting out a new definition of "because of," it failed to amend the ADEA in the same way.[53] Therefore, the Court reasoned, Congress must have intended to leave the ADEA's original, pre-1991 definition of "because of" intact.

However, the limited amendment argument cannot answer the question posed in *Gross*.[54] Recall that none of the parties in *Gross* had argued for total unification; that is, none had argued that the Court should apply the 1991 Act definition of "because of" to the ADEA. The only issue in *Gross* was partial unification; that is, whether "because of" had the same meaning in the ADEA as it did in Title VII prior to the 1991 Act (the meaning set out in *Price Waterhouse*). The limited amendment argument, which is an argument about Congress's intent in 1991, has no bearing on what the ADEA meant prior to 1991. It does nothing to rebut the presumption that when Congress used the phrase "because of" in the ADEA in 1967, it intended that phrase to have the same meaning as it did in Title VII in 1964—the meaning determined by *Price Waterhouse*.[55]

53. *See Gross*, 129 S. Ct. at 2349 ("Congress neglected to add such a provision [defining "because of"] to the ADEA when it amended Title VII to add [such a provision], even though it contemporaneously amended the ADEA in several ways"). It is worth noting that, while this argument may make sense with respect to many pre-1991 statutes, such as the ADEA (a point I contest elsewhere, *see infra* note 545), it does not seem to work for the Americans with Disabilities Act (ADA), as Congress expressly incorporated the remedy section of Title VII in that Act, including the portion of Title VII which was amended by the 1991 Act. *See* Katz, *supra* note 25, at 671 n.104.

54. In an earlier article, I argue that the limited amendment argument is flawed in five respects; that it represented a flawed view of Congress's intent in 1991. *See* Katz, *supra* note 25, at 674. My point here is that, even if we accept the limited amendment argument, it cannot do the work *Gross* needs it to do.

55. Of course, it is possible that *Price Waterhouse* got the 1964 Title VII definition wrong. But this would not help the *Gross* Court's plight. To understand this, suppose that the 1991 Act was essentially Congress's way of telling the Court, "Your attempt to define 'because of' in our 1964 statute (Title VII) was misguided." The question then becomes what Congress intended to do with respect to the 1967 statute (the ADEA). If Congress intended to correct the definition in the 1967 statute (the ADEA), as well, then the 1991 Act definition would apply in ADEA cases (total unification). Yet, this is precisely the position rejected by *Gross* in its limited amendment argument. That argument posits that Congress overruled *Price Waterhouse* only in post-1991 Title VII, leaving in place in the ADEA whatever prior definition applied: either the *Price Waterhouse* definition (partial unification) or some other definition (no unification). The limited amendment argument does not select between those two options.

Moreover, it seems highly unlikely that the 1991 Congress intended to overturn the part of *Price Waterhouse* that was in question in *Gross*—i.e., the part of *Price Waterhouse* that provided for burden-shifting upon a showing of motivating factor causation. This was precisely the part of *Price Waterhouse* that Congress incorporated into the 1991 Act.

So the *Gross* Court rejected a perfectly reasonable and widely applied canon of construction—the presumption of uniformity—with no good reason for doing so.[56]

III. ABOUT-FACE ON UNIFICATION: A RECALCITRANT COURT?

The Court has changed its view of unification. For many years, the Court embraced unification. However, beginning in 1991, the Court's enthusiasm for unification seemed to wane. And in *Gross*, the Court seems to close the door on unification. What are we to make of this timing? This Part will set out the timing of the Court's relationship with unification, and then explore potential explanations for the Court's about-face on unification. It will show that there is some evidence to support an argument that the Court's about-face represents recalcitrance toward Congress's overruling of *Price Waterhouse*. However, this Part will conclude that a better explanation for the Court's about-face on unification is its resistance to burden-shifting.

A. The Rise and Fall of Uniformity

Disparate treatment norms appear in many statutes and constitutional provisions. These norms preclude decision-makers from treating individuals adversely because of certain protected characteristics. Which characteristics are protected vary under each of these laws. For example, Title VII and the Fourteenth Amendment preclude adverse decision-making based on race or sex, labor law statutes preclude adverse decision-making based on union status, and the First Amendment precludes adverse decision-making based on the viewpoint expressed in one's speech. The question is to what extent these various statutes and constitutional provisions utilize similar standards—to what extent they are treated as being uniform.

From 1983 to 1991, the Court was all about unification. During that time, the Court consciously tried to unify standards in cases under all of the major disparate treatment laws.

For example, in 1985, in *TWA v. Thurston*,[57] the Court took a unifying position in evaluating the ADEA and Title VII. *Thurston* was

56. Although *Gross* did not offer a reason for rejecting uniformity, it did offer a reason for its choice of a simple but-for standard: that the text of the ADEA says nothing about burden-shifting. *See Gross*, 129 S. Ct. at 2348. However, this argument does not address the unification question. In 1989, the text of Title VII said nothing about burden-shifting. Yet in that year *Price Waterhouse* interpreted that same barren text as providing for burden-shifting upon a showing of motivating factor causation. The unification issue in *Gross* was whether to adopt the Title VII definition (*Price Waterhouse*) or reject that definition and look at the ADEA without reference to Title VII. The textual argument advanced by the Court in *Gross* does not answer this question.

an ADEA case dealing with discrimination in employment benefits. However, to decide *Thurston*, the Court relied on Title VII cases.[58] The Court justified this reliance on Title VII precedent by noting that the phrase "because of" in the ADEA was taken "*in haec verba*" from Title VII.[59]

Similarly, in 1987, the Court adopted a unifying stance between Title VII and the National Labor Relations Act (NLRA).[60] In *Goodman v. Lukens Steel Co.*,[61] a Title VII case against a labor union, the Court was careful to adopt an interpretation of Title VII that would be consistent with the NLRA. Again, the Court's rationale: The operative language of Title VII was taken *in haec verba* from the NLRA.[62]

And in 1989, in *Patterson v. McLean Credit Union*,[63] the Court—without even seeing a need to explain itself—applied the *McDonnell Douglas* framework from Title VII to a claim under 42 U.S.C. § 1981.[64]

But the Court's strongest unifying impulse seemed focused on the causal phrase, "because of." Specifically, the Court seemed intent on spreading a burden-shifting definition of that phrase (motivating factor causation for burden-shifting and but-for causation for liability) to as many areas of disparate treatment law as possible.

The burden-shifting definition seemed to originate in *Corning Glass Works v. Brennan*,[65] a case under the Equal Pay Act.[66] The Court then applied that definition in a pair of constitutional cases: *Mt. Healthy v. Doyle*,[67] a First Amendment speech case, and *Arlington Heights v. Metropolitan Housing Development Corp.*,[68] an Equal Protection case dealing with race discrimination.

In 1983 in *NLRB v. Transportation Management Corp.*,[69] the Court sought to determine the meaning of the phrase "because of" in the NLRA. The Court adopted a but-for standard for liability, with a

57. TWA v. Thurston, 469 U.S. 111 (1985).
58. *See id*. at 121.
59. *See id*. at 121.
60. 29 U.S.C. § 151 et seq.
61. Goodman v. Lukens Steel Co., 482 U.S. 656 (1987).
62. *Id.* at 688.
63. Patterson v. McLean Credit Union, 491 U.S. 164 (1989).
64. *See id.* at 186.
65. *Corning Glass Works v. Brennan*, 417 U.S. 188, 196 (1974).
66. Fair Labor Standards Act of 1938, 29 U.S.C. §§ 201-219 (1998).
67. Mt. Healthy City Sch. Dist. Bd. Of Educ. v. Doyle, 429 U.S. 274 (1977).
68. Vill. of Arlington Heights v. Metro. Hous. Dev. Corp., 429 U.S. 252, 270-271, n.21 (1977)
69. Nat'l Labor Relations Bd. v. Transportation Mgmt. Corp., 462 U.S. 393 (1983).

motivating factor standard for burden-shifting.[70] Notably, the Court self-consciously borrowed this standard from *Mt. Healthy*.[71]

When it was time to address the meaning of this phrase in Title VII in *Price Waterhouse*, the Court adopted the same standard: a but-for standard for liability, with a motivating factor standard for burden-shifting (at least in cases with "direct evidence").[72] Notably, all six concurring Justices sought to justify this burden-shifting definition by reference to *Mt. Healthy* and *Transportation Management*.[73] The three dissenting Justices, led by Justice Kennedy, tried to resist unification; they argued that the *Mt. Healthy* and *Transportation Management* standards should be limited to the First Amendment and NLRA, respectively. And they feared (correctly) that unification would result in the application of the *Price Waterhouse* standard to other disparate treatment statutes, such as 42 U.S.C. § 1981 and the ADEA.[74]

Between 1989 (when *Price Waterhouse* adopted its definition of "because of" in Title VII based on a unification norm) and 1991 (when Congress amended Title VII), the Supreme Court did not have occasion to preach unification. But the lower courts certainly took the Court's lead on unification. Virtually all of the lower courts to address the issue applied *Price Waterhouse*'s definition of "because of" to other disparate treatment statutes, such as the ADEA.[75]

In 1991, the tide of unification seemed to shift. In that year, Congress passed the Civil Rights Act of 1991, which partially overruled

70. *See id.* at 403.

71. *See id.* at 404 (citing *Mt. Healthy v. Doyle*, 429 U.S. 274 (1977)).

72. *See Price Waterhouse*, 490 U.S. at 228 (plurality opinion).

73. *See id.* at 248-9 (plurality opinion); *id.* at 258 (White, J., concurring); *id*. at 277 (O'Connor, J., concurring).

74. *See id.* at 292 (Kennedy, J., dissenting).

75. *See, e.g.*, Beshears v. Asbill, 930 F.2d 1348, 1353 n.6 (8th Cir. 1991) (applying *Price Waterhouse* to ADEA case); Visser v. Packer Engineering Associates, Inc., 924 F.2d 655, 662 (7th Cir. 1991) (same); Gagne v. Northwestern Insurance Company, 881 F.2d 309, 315-16 (6th Cir. 1990) (same); *see also* Wilson v. Firestone Tire & Rubber Co., 932 F.2d 510, 514 (6th Cir. 1991) (declining to apply *Price Waterhouse*'s burden-shifting mechanism in ADEA case where there was no "direct evidence"; in other words, applying *Price Waterhouse*'s "direct evidence" distinction in ADEA case). Courts also routinely applied *Price Waterhouse* to other disparate treatment statutes other than Title VII or the ADEA. *See, e.g.*, Randle v. LaSalle Telecommunications, Inc., 876 F.2d. 563 (7th Cir. 1989) (applying *Price Waterhouse* framework to claim under 42 U.S.C. §1981); *see also* Sischo-Nownejad v. Merced Cmty. College Dist., 934 F.2d 1104 (9th Cir. 1991) (applying Title VII analysis to claim under 42 U.S.C. §1983); Merrick v. Farmers Insurance Group, 892 F.2d 1434 (9th Cir. 1990) (applying Title VII analysis to claim of retaliation under ADEA). The only court of appeals during this era that seemed to depart from this assumption of uniformity did so as one of two alternative grounds in an unpublished decision. *But see* Narang v. Chrysler Corp., 896 F.2d 1369 (Table), 1990 WL 18057, at *4 n.2 (6th Cir. 1990) (declining to apply *Price Waterhouse* in case brought under 42 U.S.C. § 1981).

Price Waterhouse (and a number of other Supreme Court cases). While both Congress and the Court agreed that burden-shifting was appropriate upon a showing of motivating factor causation, they disagreed on the standard for liability. *Price Waterhouse* had defined "because of" as requiring but-for causation for liability. Congress saw this definition as overly restrictive, and therefore amended Title VII in the 1991 Act to make clear that "because of" required only motivating factor causation for liability (with the but-for standard determining only the availability of full damages).[76]

After this, the courts began to lose their enthusiasm for unification. The anti-unification movement began in the lower courts, most of which rejected the idea that Congress's 1991 Act definition would apply to disparate treatment statutes other than Title VII.[77] That is, the lower courts almost immediately rejected total unification. However, most of these courts still applied *Price Waterhouse* in non-1991 Act cases.[78] That is, they still tended to accept partial unification.

Initially, the Supreme Court appeared to be on the fence. On one hand, the Court appeared to endorse the lower courts' rejection of total unification. The Court routinely denied certiorari in lower court cases that had rejected total unification.[79] Moreover, in dicta in a 2005 disparate impact case, the Court seemed to recite a version of the limited

76. *See* Mardell v. Harleysville Life Ins. Co., 31 F.3d 1221, 1235 n.23 (3d Cir. 1994) ("One overriding lesson the 1991 Act tutors . . . is that Congress was unhappy with increasingly parsimonious constructions of Title VII."); EEOC Policy Guidance No. 915.002, ¶ 2095 n.14 (July 14, 1992); 2 LARSON, EMPLOYMENT DISCRIMINATION § 35.04[1] (2009). *See also* Civil Rights Act of 1991, Pub. L. No. 102-166, § 3, 105 Stat. 1071, 1071 (codified as amendment at 42 U.S.C. § 1981) (stating that the purpose of the Act was "to respond to recent decisions of the Supreme Court by expanding the scope of relevant civil rights statutes in order to provide adequate protection to victims of discrimination").

77. *See*. Katz, *supra* note 25, at 647 n.22 (2008) (listing lower court cases rejecting total unification).

78. *See, e.g.*, Equal Employment Opportunity Comm'n v. Warfield-Rohr Casket Co., 364 F.3d 160 (4th Cir. 2004) (applying *Price Waterhouse* framework to ADEA claim); Febres v. Challenger Caribbean Corp, 214 F.3d 57 (1st Cir. 2000) (same); Miller v. Cigna Corp., 47 F.3d 586 (3rd Cir. 1995) (same): Ostrowski v. Atlantic Mutual Insurance Co, 968 F.2d 171 (2d. Cir. 1992) (same); *see also* Kubicko v. Ogden Logistics Servs., 181 F.3d 544 (4th Cir. 1999) (applying *Price Waterhouse* framework to a Title VII retaliation claim); Kelly v. Drexel University, 907 F.Supp. 864, 870-71 (E.D. Pa. 1995) (applying *Price Waterhouse* to ADA case); Hutchinson v. United Parcel Serv., Inc., 883 F.Supp. 379, 399 (N.D. Iowa 1995) (same); Reiff v. Interim Personnel, Inc., 906 F.Supp. 1280, 1286 (D.Minn. 1995) (declining to apply *Price Waterhouse*'s burden-shifting mechanism in ADA case where there was no "direct evidence"; in other words, applying *Price Waterhouse*'s "direct evidence" distinction in ADA case).

79. *See* Katz, *supra* note 25, at 647 n.22 (listing cases denying certiorari).

amendment argument—the primary argument advanced by lower courts for rejecting total unification in disparate treatment cases.[80]

On the other hand, in four post-1991 cases, the Court arguably endorsed at least some version of unification. In *St. Mary's Honor Center v. Hicks*,[81] *Hazen Paper Co. v. Biggins*,[82] *Reeves v. Sanderson Plumbing*,[83] and *Raytheon Co. v. Hernandez*,[84] the Court was willing to assume that the *McDonnell Douglas* framework, developed in a Title VII case, applied in cases brought under other statutes (Section 1981, the ADEA, and the ADA). The Court's willingness to assume the application of *McDonnell Douglas* in these four cases might suggest some continued support for the concept of unification.

But in 2009, the Court decisively rejected all forms of unification in *Gross*. The Court expressly rejected total unification, refusing to apply the 1991 Act definition to a case under the ADEA.[85] And the Court also rejected partial unification, refusing to apply the *Price Waterhouse* definition of "because of" to the ADEA.[86] The age of unification was officially at an end.

80. *See* Smith v. City of Jackson, 544 U.S. 228 (2005). In *Smith*, the Court held that disparate impact claims are allowable under the ADEA. In discussing what standards might apply to disparate impact claims under the ADEA, the Court noted that the 1991 Act amended only Title VII's disparate impact provisions—thereby suggesting that pre-1991 Act law might apply to non-Title VII disparate impact provisions. *See id*. at 240. This is the limited amendment argument, discussed above. *See supra* note 53 and accompanying text. The lower courts that rejected total unification tended to use this same argument. *See* Katz, *supra* note 25, at 647 n.22.

The Court arguably rejected the presumption of uniformity in General Dynamics v. Cline, 540 U.S. 581 (2004). In that case, the Court held that, unlike Title VII, which had been interpreted as proscribing "reverse discrimination" (discrimination against non-minorities), the ADEA did not proscribe discrimination against younger employees. *See id.* at 590. However, this holding interpreted the word "age," which is unique to the ADEA—not the phrase "because of," which is common to the two statutes. Moreover, the holding turned on an argument about legislative history and intent of the ADEA (an intent to protect older workers, not younger workers), which the Court argued was different from the history and intent of Title VII. *See id*. Therefore, *General Dynamics* seems to fit within one of the exceptions to the uniformity rule, *see supra* notes 50 and 51, as opposed to representing a clear rejection of that rule.

81. St. Mary's Honor Ctr. v. Hicks, 509 U.S. 502, 506 n.1 (1993) (discussing 42 U.S.C. § 1981).

82. Hazen Paper Co. v. Biggins, 507 U.S. 604, 612 (1993) (discussing ADEA).

83. Reeves v. Sanderson Plumbing Prods., Inc, 530 U.S. 133, 141 (2000) (discussing ADEA).

84. Raytheon Co. v. Hernandez, 540 U.S. 44 (2003) (discussing ADA).

85. *See* Gross v. FBL Fin Servs., 129 S. Ct. 2343, 2349 (2009).

86. *See id*.

B. An Act of Recalcitrance?

The timing of the Court's about-face on uniformity might suggest petty motives. After all, when unification involved expanding the application of the Court's definition of "because of," the Court enthusiastically embraced unification. But after 1991, when unification would involve expanding the application of Congress's definition—a definition developed as a rebuke to the Court—the Court seemed to lose its enthusiasm for unification. This timing might suggest that the Court's rejection of the unification doctrine might contain an element of recalcitrance; antipathy either to being rebuked by Congress or to the specific definition imposed by Congress.[87]

However, as intuitively appealing as this timing argument might be, there are two potential flaws in it. First, the Court's post-1991 rejection of total unification may have been justified. There is a potential justification—albeit not one directly advanced by the Court—for declining to apply a presumption of uniformity in cases involving subsequent amendment, like the 1991 Act.

The doctrine of unification is arguably based on the timing. At Time 1, Congress uses a phrase (such as "because of") in Statute 1 (here, Title VII). Later, at Time 2, Congress uses the same phrase in Statute 2 (here, the ADEA). The presumption of uniformity is based on the idea that, at Time 2, Congress intended to adopt the meaning of the phrase used in Statute 1. In other words, at Time 2, Congress intended to look back to something it had done before (at Time 1) and incorporate it into what it was doing at Time 2.

This concept encounters problems when Statute 1 is subsequently amended at Time 3 (here, in the 1991 Act). At Time 2, one can imagine Congress seeking to incorporate a meaning developed at Time 1. But it is harder to imagine that, at Time 2, Congress intended to incorporate a meaning that it would develop sometime in the future at Time 3.

A variation of this timing problem was advanced by the *Gross* Court: It is possible that Congress's intent at Time 3 (1991) trumps any intention it may have had at Time 2 (1967). That is, irrespective of any

87. Some authors prior to *Gross* had suggested that the federal courts might be biased against employment law plaintiffs. *See, e.g.*, Kevin M. Clermont et al., *How Employment-Discrimination Plaintiffs Fare in the Federal Courts of Appeals*, 7 EMP. RTS. & EMP. POL'Y J. 547 (2004) (noting discrepancies between plaintiffs' and defendants' success at the appellate level); Kevin M. Clermont & Stewart J. Schwab, *How Employment Discrimination Plaintiffs Fare in Federal Court*, 1 J. EMPIRICAL LEGAL STUD. 429 (2004); Kevin M. Clermont & Stewart J. Schwab, *Employment Discrimination Plaintiffs in Federal Court: From Bad to Worse?*, 3 HARV. L. & POL'Y REV. 103 (2009). These authors might believe that *Gross*'s rejection of the uniformity norm reflects anti-plaintiff bias. This issue is beyond the scope of this Article.

intent Congress might have had in 1967 to tether the ADEA to Title VII—either as it existed in 1964 or even as it might be amended in the future—the 1991 Congress may have intended to untether those two statutes. The limited amendment argument made by the *Gross* Court posits that Congress's decision to amend Title VII without amending the ADEA represented a decision to de-couple the meaning of "because of" in those two statutes.

The point of these arguments about timing is not that they should have carried the day in *Gross* or that Congress's amendment of Title VII in the 1991 Act precluded total unification.[88] Rather, the point is that a reasonable argument could have been made for the rejection of total unification after 1991. The existence of such arguments might give pause to those who would see Court's post-1991 rejection of total unification as an act of recalcitrance.[89]

88. In fact, there are at least two arguments that might overcome this timing issue—two ways the presumption of uniformity might apply even in cases involving subsequent amendment. First, in Statute 2, Congress might have intended to adopt a dynamic, rather than a static, meaning of the phrase it used in Statute 1 ("because of"). In other words, Congress might have intended to adopt the meaning of the phrase as it might later be amended (dynamic), rather than the meaning of that phrase as of Time 1 (static). *Gross* does not appear to have contemplated this argument.

Second, irrespective of whether Congress intended to adopt a static or dynamic meaning at Time 2, uniformity would be appropriate if Congress intended to impose uniformity at Time 3. In other words, at Time 3, when Congress amended the definition of Statute 1's phrase "because of" in Statute 3, Congress might intend its new definition to apply to all existing statutes using the same phrase. *Gross* advanced the limited amendment argument, which effectively rejects the idea that Congress intended to impose uniformity at Time 3 (1991). The limited amendment argument assumes that Congress's failure to amend other statutes in 1991 indicates an intent not to apply its 1991 Act definition of "because of" to those statutes. (I have argued elsewhere that the limited amendment argument cannot bear this weight). *See* Katz, *supra* note 25, at 674.

89. In fact, one might argue that *Gross* put an end to a form of recalcitrance in the lower courts. As noted above, prior to *Gross*, many lower courts continued to apply *Price Waterhouse* after that case had been overruled. They used the doctrine of partial unification to apply a Title VII case that had been overruled by an amendment of Title VII to non-Title VII cases. *Gross*'s rejection of partial unification avoids this problem. However, this view would be overly optimistic. First of all, there is nothing in *Gross* to indicate that it was concerned with continuing to apply a Congressionally overruled precedent. Second, by rejecting the doctrine of partial unification and overruling *Price Waterhouse*, *Gross* actually wiped out the one part of *Price Waterhouse* that Congress had agreed with: burden-shifting.

A similar form of lower court recalcitrance might be seen in the fact that prior to the Supreme Court's intervention in *Desert Palace*, all but one circuit had attempted to limit the application of the 1991 Act by imposing *Price Waterhouse*'*s* "direct evidence" requirement. *See, e.g.*, Mohr v. Dustrol, Inc., 306 F.3d 636, 640-641 (8th Cir. 2002); Fernandes v. Costa Bros. Masonry, Inc., 199 F.3d 572, 580 (1st Cir. 1999); Trotter v. Bd. of Trustees of Univ. of Ala., 91 F.3d 1449, 1453-1454 (11th Cir. 1996); Fuller v. Phipps, 67 F.3d 1137, 1142 (4th Cir. 1995). The Supreme Court put a stop to this in Desert Palace v. Costa, 539 U.S. 90 (2003). But again, there is nothing in *Desert Palace* to suggest that the Court was concerned with lower court recalcitrance—that is, with the

But there is a second, and perhaps more serious, problem with the recalcitrance theory: *Gross* went too far to support such a theory. Recalcitrance might arguably explain the Court's rejection of total unification—that is, its refusal to apply Congress's new 1991 standard in non-Title VII cases. However, in *Gross*, the Court went a step further than this, rejecting partial unification—that is, rejecting the application of its own *Price Waterhouse* standard in such cases. This looks more like the act of a Court that is having second thoughts about its own definition of "because of" than the act of a Court that is petulant because its definition was rebuked by Congress.[90]

C. An Alternative View: Resistance to Burden-Shifting

A better explanation for the Court's about-face on unification may be that the Court changed its view of burden-shifting as part of the definition of "because of." Recall that, prior to 1989, the Court seemed to use unification as a vehicle to spread burden-shifting (where a plaintiff proves motivating factor causation) to numerous disparate treatment doctrines. The burden-shifting doctrine started in the Equal Pay Act, and spread to the Equal Protection Clause and the First Amendment of the Constitution, as well as to the NLRA. In *Price Waterhouse*, the Court unified Title VII with those other areas, holding that Title VII's phrase "because of" permits a burden-shift upon a showing of motivating factor causation.

Yet, from the moment *Price Waterhouse* was decided, burden-shifting was controversial. Three Justices dissented in *Price Waterhouse*, arguing against unification and against burden-shifting. In addition, Justice O'Connor concurred to make clear that she considered burden-shifting to be "strong medicine," which should be limited to cases where the plaintiff could show "direct evidence."[91] For many years, the circuit courts struggled with the meaning of "direct evidence," splitting four ways over the meaning of that phrase—and therefore over when plaintiffs could use the burden-shifting doctrine.[92]

After 1991, the Court could not prevent burden-shifting in 1991 Act Title VII cases. Congress had mandated burden-shifting in those cases. But for 12 years, many of the lower courts tried to limit burden-shifting

fact that the lower courts seemed grudging in their willingness to apply Congress's new statute.

90. The lower courts had not gone this far. They generally rejected full unification with the 1991 Act, but did not reject partial unification (that is, the application of the earlier judicial definition of "because of" in *Price Waterhouse*). *See* Katz, *supra* note 25, at 647 n.22.

91. *See Price Waterhouse*, 490 U.S. at 262 (O'Connor, J., concurring).

92. *See Desert Palace*, 539 U.S. 90.

even in 1991 Act cases, applying the "direct evidence" doctrine from the pre-1991 case of *Price Waterhouse* to limit the availability of burden-shifting to a sub-set of 1991 Act cases.[93] Although the Supreme Court put a stop to this practice in *Desert Palace* in 2003, it did so on narrow textual grounds.[94] The Court's decision in *Desert Palace* did not say a word in support of burden-shifting.

The Court's coolness toward burden-shifting may also have been evident in its limited embrace of uniformity after *Price Waterhouse*. After that, the Court addressed uniformity only four times in disparate treatment cases: In *St. Mary's Honor Center. v. Hicks*,[95] *Hazen Paper Co. v. Biggins*,[96] *Reeves v. Sanderson Plumbing Products, Inc.*,[97] and *Raytheon Co. v. Hernandez*,[98] the Court was willing to assume uniformity; to assume (without so holding) that Title VII doctrine applied in other disparate treatment statutes. But notably, in all four of these cases, the Title VII doctrine in question was the *McDonnell Douglas* simple but-for test—not the burden-shifting mechanism of *Price Waterhouse.* In other words, the Court seemed to be comfortable accepting the possibility of unification in cases that did not involve burden-shifting; comfortable where unification did not expand the use of burden-shifting. But the Court had gone silent on unification where it would expand burden-shifting.

Finally, in 2009, the *Gross* Court seemed to find the votes to rid itself of burden-shifting in non-1991 Act cases—to effectively overrule *Price Waterhouse*.[99] Notably, two of those votes came from Justice Kennedy, who had written the dissent in *Price Waterhouse,* and Justice Scalia, who had joined that dissent.

What is important to note is the relationship that developed between burden-shifting and uniformity. Burden-shifting had been brought into Title VII as a result of *Price Waterhouse*'s insistence on uniformity between Title VII and other constitutional and statutory disparate

93. *See id.*

94. *See id.* (holding that Congress's failure to use the phrase "direct evidence" in the 1991 Act suggested a lack of intent to make this distinction).

95. St. Mary's Honor Ctr. v. Hicks, 509 U.S. 502, 506 n.1 (1993) (discussing 42 U.S.C. § 1981).

96. Hazen Paper Co. v. Biggins, 507 U.S. 604, 612 (1993) (discussing ADEA).

97. Reeves v. Sanderson Plumbing Prod., Inc., 530 U.S. 133, 141 (2000) (discussing ADEA).

98. Raytheon Co. v. Hernandez, 540 U.S. 44, 49-50 (2003) (discussing ADA).

99. *See* Hart, *supra* note 41 (forthcoming 2009) (manuscript on file with author) ("[W]hile the Court's decision never explicitly overrules Price Waterhouse, its reasoning directly contradicts the reasoning of that opinion and it is hard to imagine what is left of Price Waterhouse after Gross. Most significantly, the Gross majority adopts the causation standard pressed unsuccessfully by the Price Waterhouse dissenters.").

treatment doctrines.[100] Lower courts had used the doctrine of uniformity to expand *Price Waterhouse*'s burden-shifting mechanism to other disparate treatment statutes, such as the ADEA.[101] These cases were premised on the idea that, when Congress used the phrase "because of" in the ADEA, it meant to adopt the meaning of that phrase from the 1964 version of Title VII. And *Price Waterhouse* had defined that phrase in 1964 Title VII as including burden-shifting upon a showing of motivating factor causation.[102]

So to reject burden-shifting in the ADEA, the Court needed to untangle that statute not just from the 1991 Act (which it did through the limited amendment argument), but also from pre-1991 Title VII. That is, the Court needed to reject not only total unification; it also needed to reject partial unification, as it did in *Gross*.[103]

IV. *GROSS*'S PROBLEMATIC DEFINITION OF "BECAUSE OF"

The Court's rejection of burden-shifting is normatively problematic. This Part will explain why burden-shifting is desirable in disparate treatment law. It will also expose three additional normative flaws in *Gross*'s definition of "because of": *Gross*'s definition lets discriminators get away with discrimination, under-deters discrimination, and unfairly allocates windfall in over-determined cases entirely to defendants. Accordingly, this Part will conclude, Congress should overrule *Gross*.[104]

100. *See Price Waterhouse*, 490 U.S. at 248.

101. *See* Katz, *supra* note 25, at 647 n.22.

102. *See supra* notes 20-22 and accompanying text.

103. *See* Gross v. FBL Fin. Serv. Inc., 129 S. Ct. 2343, 2352 n.6 (2009) (rejecting application of *Transportation Management* or *Mt. Healthy* beyond the NLRA and First Amendment, respectively). Interestingly, it is not clear that the *Gross* Court had the courage of its convictions in this regard. If the Court had faith in the persuasiveness of its rejection partial unification, it could have stopped after its first argument—that there was no mention of burden-shifting in the ADEA. After all, if the ADEA truly stood alone, with no connection to pre-1991 Title VII, that argument would have sufficed. But the Court felt compelled to go on to overrule *Price Waterhouse*, a pre-1991 Title VII case. This might suggest that, despite rejecting unification, the Court continued to feel the pull of its reasoning. That is, the Court might have worried that readers would continue to be drawn to pre-1991 Title VII to interpret the ADEA, and therefore felt compelled to attack burden-shifting in pre-1991 Title VII, as well as the ADEA.

104. In an earlier article, I had argued that the courts could avoid the normative flaws discussed in this Part without intervention from Congress. *See* Katz, *supra* note 25, at 643. Several other scholars have disagreed. *See, e.g.*, Corbett, *supra* note 1, at 83 (2009) (arguing that legislative action is required to unify disparate treatment law); Jamie Darin Prenkert, *The Role of Second-Order Uniformity in Disparate Treatment Law:* McDonnell Douglas*'s Longevity and the Mixed-Motives Mess*, 45 AM. BUS. L.J. 511 (2008) (same); Steven J. Kaminshine, *Disparate Treatment as a Theory of Discrimination: The Need for a Restatement, Not A Revolution*, 2 STAN. J. CIV. RTS. CIV. LIBERTIES 1, 7 (2005) (same); *see also* Howard Eglit, *The Age Discrimination in Employment Act, Title VII, and the Civil Rights Act of 1991: Three Acts and a Dog that Didn't Bark*, 39 WAYNE L. REV.

A. The Benefits of Burden-Shifting

As discussed above, *Gross* adopted a but-for standard of causation with no possibility of burden-shifting. Such a standard is normatively problematic, as it makes proving but-for causation unduly difficult for plaintiffs.

Proving any form of causation is difficult for disparate treatment plaintiffs. Causation occurs in the mind of the decision-maker/defendant. And most of the relevant evidence tends to be under the control of the defendant. This lack of access to evidence makes proving any type of causation difficult, and therefore makes burden-shifting normatively desirable.[105]

But proving but-for causation is particularly difficult. What keeps a factor from being a but-for cause is the existence of a second, independently sufficient factor. To understand this, consider the ubiquitous two-fires hypothetical from first-year Torts class: Suppose that D starts a fire, which engulfs P's house and burns it down. If there is no other factor that would have burned down P's house, D's fire will be a but-for cause of the house burning down. But now suppose that there was a second fire, say one started by a lightning strike. And suppose that the second fire reaches the house at the same time as D's fire, and that the second fire would have burned the house down irrespective of D's fire. In such a case, the second, independently sufficient factor (the lightning fire) prevents the first fire (D's fire) from being a but-for cause.[106]

The same concept applies in disparate treatment cases. Suppose that D fires P. And suppose that one of the factors in D's decision was P's sex. If there are no other factors that would have triggered D's decision to fire P, then P's sex would be a but-for cause of that decision.

1093, 1215 (1993) (arguing that only way to apply a 1991 Act standard to the ADEA would be through an act of Congress or the Supreme Court). While I still believe that, before *Gross*, disparate treatment law could have been unified without legislative action, after *Gross*, it now appears that legislative action is required.

105. *See* CHRISTOPHER B. MUELLER & LAIRD C. KIRKPATRICK, EVIDENCE 105 (3d ed. 2003) (it is often appropriate to shift the burden of proof to the party who has the greatest access to evidence). *See also Gross*, 129 S. Ct. at 2359 (Breyer, J., dissenting) ("[S]ince the employee likely knows less than does the employer about what the employer was thinking at the time, the employer will often be in a stronger position than the employee to provide the answer.").

106. This hypothetical is probably based on the case of Cook v. Minneapolis, St. P. & S. S. M. Ry. Co., 74 N.W. 561 (Wis. 1898). However, most textbooks raising this hypothetical tend to use the later case of Kingston v. Chicago & N.W. Ry. Co., 211 N.W. 913 (Wis. 1927) (citing and distinguishing *Cook*). *See, e.g.*, JAMES HENDERSON, RICHARD PEARSON & JOHN SICICIANO, THE TORTS PROCESS 145-47 (5th Ed. 1999). The similarity between disparate treatment law and the common law of torts has often been noted. *See, e.g.*, *Price Waterhouse*, 490 U.S. at 264 (O'Connor, J., concurring).

But suppose that D also considered another factor in his decision, such as P's chronic tardiness. And suppose that D would have fired P for chronic tardiness irrespective of her sex. In such a case, the second, independently sufficient factor (chronic tardiness) prevents the first factor (sex) from being a but-for cause.

What this means is that, to prove but-for causation, a plaintiff essentially has to prove a negative: the absence of any other independently sufficient cause. However difficult this may be in the physical world of torts, where there are a limited number of observable causes for an event, it tends to be extremely difficult in anti-discrimination law, where decision-makers may have any number of reasons for a decision, none of which are observable and any of which might be independently sufficient to reach the decision.[107]

One might hope that modern discovery would alleviate this problem somewhat. If the plaintiff survives a motion for judgment on the pleadings, her lawyer will likely get to depose the decision-maker. In the deposition, the lawyer will likely ask the time-honored two questions: (1) what reasons did you have for making the decision you made, and (2) are there any other reasons? Thus, an effective plaintiff's lawyer might at least narrow the universe of potential independently sufficient reasons.

But many plaintiffs do not get past motions to dismiss, particularly after *Ashcroft v. Iqbal*, which may raise pleading standards in employment discrimination cases.[108] Moreover, even if a plaintiff survives a motion to dismiss, deposes the decision-maker, and asks the "any other reason" question, the plaintiff will likely have great difficulty proving but-for causation. This is because, to prove but-for causation, the plaintiff has to disprove every non-discriminatory reason advanced by the decision-maker[109] (and occasionally some that the court might posit, though this practice by judges is improper).[110]

Burden-shifting provides a good solution for the problems inherent in proving but-for causation. Under burden-shifting as contemplated by *Price Waterhouse* or the 1991 Act, the plaintiff is required to prove motivating factor causation in order to shift the burden. This requirement avoids the problem of treating a defendant as guilty until

107. *See Gross*, 129 S. Ct. at 2358 (Breyer, J., dissenting) ("In [tort law], reasonably objective scientific or commonsense theories of physical causation make the concept of 'but-for' causation comparatively easy to understand and relatively easy to apply. But it is an entirely different matter to determine a 'but-for' relation when we consider, not physical forces, but the mind-related characterizations that constitute motive.").

108. Ashcroft v. Iqbal, 129 S. Ct. 1937 (2009).

109. *See* Katz, *supra* note 1, at 139.

110. *See id*. at 170.

proven innocent: The burden does not fall on the defendant until it is proven that the defendant engaged in wrongdoing (consideration of a protected factor in employment decision-making). These burden-shifting schemes then place the burden on the defendant—the party that (1) caused the uncertainty in the first place (by considering a protected factor), (2) has the best access to evidence, and (3) has the best information about potential non-discriminatory reasons for its decision.[111]

These considerations led the Court in *Price Waterhouse* and then Congress in the 1991 Act to adopt such a burden-shifting standard.[112] But *Gross* departs from these principles, requiring the plaintiff to shoulder the entire burden without regard to the difficulty of proving but-for causation, the fact that most of the evidence on the issue is under the control of the defendant, or the fact that the defendant has engaged in wrongdoing. The *Gross* standard is therefore normatively problematic in this regard.

Gross defends its rejection of burden-shifting with a rather summary assertion that *Price Waterhouse*'s "burden-shifting framework is difficult to apply."[113] However, the difficulty with *Price Waterhouse*'s burden-shifting framework has not been the framework itself. The framework simply asks two questions: did the plaintiff prove that the protected characteristic (e.g., sex) was a motivating factor, and if so did the defendant prove that it would have reached the same decision absent that characteristic. There is nothing about either of these two inquiries that is particularly difficult.

What has been difficult about the burden-shifting framework has been figuring out when to apply it. Specifically, the courts have repeatedly tried to place limits on the applicability of that framework, many of which have proven unworkable. The most notable of these limits has been Justice O'Connor's "direct evidence" limit,[114] which caused a four-way split among the Courts of Appeals over the meaning of "direct evidence."[115] But these limits—and their ambiguity—are

111. *See* MUELLER & KIRKPATRICK, *supra* note 101, at 105 (these are factors that suggest the appropriateness of burden-shifting).

112. *See supra* notes 20-22 and accompanying text (discussing *Price Waterhouse*) and note 29 and accompanying text (discussing 1991 Act).

113. *See Gross*, 129 S. Ct. at 2352.

114. *See* Price Waterhouse v. Hopkins, 490 U.S. 228, 277 (1989) (O'Connor, J., concurring) (requiring "direct evidence" for burden-shift).

115. *See* Costa v. Desert Palace, Inc., 299 F.3d 838, 852-53 (9th Cir.), *aff'd* 539 U.S. 90 (2003) (explaining various definitions of "direct evidence"). Other courts have adopted other convoluted ways to limit burden-shifting. *See* Katz, *supra* note 1, at 119 (discussing courts' "creative" methods of trying to limit the applicability of the burden-shifting framework in the 1991 Act).

products of courts' discomfort with and consequent desire to limit burden-shifting. The solution to the difficulty noted by *Gross* is not to reject burden-shifting. Rather, the solution is to reject ambiguous and difficult limits on burden-shifting, such as the "direct evidence" requirement.[116]

In summary, burden-shifting makes good sense in the context of proving but-for causation. *Gross* rejected burden-shifting based on a reason that made little sense.

Yet the Court in *Gross* has made clear that it is hostile to burden-shifting; it went through significant lengths to avoid it, even going so far as to reject the widely accepted canon of uniformity. So we can likely expect the Court to stick to its position, and to reject burden-shifting in any statute that does not expressly require it. *Gross* rejected burden-shifting under the ADEA. It seems likely that, in the future, the Supreme Court (or lower courts applying *Gross*) will reject burden-shifting in most disparate treatment statutes other than the 1991 Act—in statutes such as the Americans with Disabilities Act, Family and Medical Leave Act, and various anti-retaliation statutes.[117] Accordingly, Congress should make clear that all of its disparate treatment statutes permit burden-shifting.

B. The Problems of a But-For Standard For Liability

There is another problematic aspect of the definition of "because of" adopted in *Gross*:[118] It requires but-for causation for liability. There are three problems with this requirement.

First, a but-for standard for liability permits some discriminators to get away with discrimination. Specifically, this standard lets some defendants who engage in motivating factor discrimination—who consider protected factors, such as race, sex, or age in their decision-making—escape liability.

Presumably, one goal of disparate treatment law is to punish those who violate the norm against considering protected factors, such as race,

116. *See* Corbett, *supra* note 5, at 25 (referring to "direct evidence" requirement as "chimerical" and noting that "it is likely that no one sheds a tear for the Court abolishing [this] bad standard" in *Desert Palace*).

117. Some lower courts have already acted based on this prediction. *See, e.g.*, Fairly v. Andrews, 578 F.3d 518 (7th Cir. 2009) (applying *Gross* to First Amendment free speech claim); Levi v. Wilts, No. 08-3042, 2009 WL 2905927 (C.D.Ill. Sept. 4, 2009) (applying *Gross* to a constitutional retaliation claim); Williams v. District of Columbia, 646 F. Supp. 2d 103 (D.D.C. 2009) (applying *Gross* to a claim brought under the Juror Act).

118. *See supra* note 40.

sex, or age in their employment decision-making.[119] In fact, the Court has stated repeatedly (and many of the Justices in the *Gross* majority have stated) that any race-based decision-making is problematic.[120] Yet, if we are serious about punishing decision-making that considers protected characteristics such as race, reserving liability for but-for causation is too limiting.

A but-for requirement for liability means that there will be cases in which a decision-maker can consider protected characteristics, such as race, sex, or age, with impunity. Under a but-for standard, as long as the decision-maker also relied on a second, independently sufficient legitimate factor (such as chronic tardiness), there is no liability. Such a decision-maker will suffer no adverse consequences, despite his consideration of a protected factor in his decision-making. Thus, a but-for standard for liability permits wrongdoers to escape punishment.

A second, related problem with *Gross's* but-for test for liability is under-deterrence. If those who consider protected characteristics can, in some cases, can get away with such discrimination, then there is less than optimal deterrence. Those who discriminate will know that they have a chance of getting away with discrimination, as long as they can find a second, independent cause for their decision. And, given the realities of most employment decision-making, there is a high probability that most defendants will be able to claim that they relied upon a second, independent (and non-discriminatory) factor.[121]

A motivating factor test for liability—such as that found in the 1991 Act[122]—avoids these problems. A motivating factor standard for liability ensures that decision-makers who consider protected characteristics, such as race, sex, or age, in their decision-making, are held accountable, thus providing deterrence.[123]

A third problem with *Gross*'s requirement of but-for causation for liability is that it unfairly allocates a windfall to defendants. In cases with more than one sufficient factor (over-determined cases), any all-or-

119. *See* Price Waterhouse v. Hopkins, 490 U.S. 228, 264-65 (1989) (O'Connor, J., concurring) (noting deterrence as goal of Title VII, and noting that, "There is no doubt that Congress considered reliance on gender or race in making employment decisions an evil in itself.").

120. *See* City of Richmond v. J.A. Croson Co., 488 U.S. 469 (1989); Parents Involved In Cmty. Sch. v. Seattle Sch. Dist., 551 U.S. 701, 748 (2007) ("[T]he way to stop discrimination on the basis of race is to stop discriminating on the basis of race.").

121. While second, independent factors may be rare in tort law, as in the two-fires hypothetical, they appear quite commonplace in anti-discrimination law.

122. *See supra* note 29.

123. In an earlier article, I question whether merely attaching liability, with limited damages, provides adequate deterrence. *See* Katz, *supra* note 14, at 534. But it at least it provides some deterrence in such cases, which is more than can be said for *Gross*'s requirement of but-for causation for liability.

none standard (such as a but-for requirement) for damages will yield a windfall to one of the parties. To understand this, recall the two-fire hypothetical. In that example, two fires, one of which was started by D, converged on P's house and burned it down. Either fire would have burned down P's house. In such a case, D will argue that he should not have to pay P for the cost of the house. If he does pay P, then P will be put in a better position than P would have been in had D not started his fire. In other words, a rule that required D to pay P would give P a windfall. For this reason, it is tempting to apply a but-for requirement for liability. Such a rule prevents P from receiving such a windfall.

However, a but-for requirement in such a case provides a windfall to D. Had there not been a second fire, D's fire would have simply burned down P's house and D would have had to pay P. The but-for requirement permits D to avoid paying P merely because of the second fire. In other words, the but-for requirement puts D in a better position than he would have been in absent the second fire; it provides a windfall to D.

Gross's but-for requirement for liability allocates windfall to defendants in much the same way—in a context in which over-determination is much more likely to occur than the proverbial second fire.[124] Suppose that a defendant considers age in its decision to demote the plaintiff (as Jack Gross alleged FBL did). And suppose that consideration of age alone would have resulted in the decision to demote. Absent a second, independent factor, the defendant would be liable and have to pay the plaintiff for the cost of his demotion. But now suppose that the defendant were presented with a second, independently sufficient reason for its decision (such as corporate restructuring, as alleged by FBL). In such a case, a but-for standard would result in no liability—in a windfall for the defendant.[125]

The *Gross* Court claimed that but-for causation was required by "ordinary usage," looking to the definitions of "because of" in three popular dictionaries.[126] But this claim is patently false. The portion of the dictionaries quoted by the court say nothing about but-for causation. Rather, the definitions highlighted by the Court merely speak about

124. *See id.*

125. The issue is slightly more complicated when the second factor results from wrongdoing by the plaintiff, as in a case where the second fire was started by the plaintiff rather than by lightning. This is often the situation in employment cases, such as where the second factor is something like the plaintiff's chronic tardiness. However, the same basic principles apply. *See* Katz, *supra* note 14, at 489.

126. *See* Gross v. FBL Fin. Serv. Inc., 129 S. Ct. 2343, 2350 (2009) (citing 1 WEBSTER'S THIRD NEW INTERNATIONAL DICTIONARY 194 (1966); 1 OXFORD ENGLISH DICTIONARY 746 (1933); THE RANDOM HOUSE DICTIONARY OF THE ENGLISH LANGUAGE 132 (1966)).

causation generally, using phrases such as "by reason of; on account of."[127] These definitions do not specify which of the many possible causal concepts they reference. "By reason of" or "on account of" might mean necessity/but-for causation, but it might just as well mean sufficiency, or minimal/motivating factor causation.[128]

To shore up its dictionary argument, the Court looks to the law of torts for support, quoting a 1984 treatise for the proposition that causation means but-for causation.[129] This is ironic, given that modern tort law has actually rejected the but-for test adopted by *Gross* in favor of a standard which better addresses the windfall problem in over-determined cases.[130]

In tort cases where an independently sufficient factor (like the second fire, or an employee's chronic tardiness) is present, and thereby precludes the defendant's wrongdoing from being a but-for cause of the plaintiff's harm, the modern courts provide an alternative avenue for liability: These courts will hold a defendant liable even if his conduct is not a but-for cause of plaintiff's harm, as long as the defendant's conduct is a "substantial factor" in causing the plaintiff's harm.[131] And while that phrase—"substantial factor"—is, by itself, ambiguous, the Restatement of Torts has recognized that this phrase includes sufficiency.[132] Thus, in cases where a defendant's conduct is sufficient, but not necessary (not a but-for cause), modern tort law imposes liability on the defendant.[133] Then, so as not to allocate all of the windfall to the plaintiff, modern tort law generally allocates damages based on relative fault.[134] (In the two-fires hypothetical, for example, if the second fire was caused by the plaintiff's negligence or by a third party, as opposed to a lightning strike,

127. *See* 1 WEBSTER'S THIRD NEW INTERNATIONAL DICTIONARY 194 (1966); 1 OXFORD ENGLISH DICTIONARY 746 (1933); THE RANDOM HOUSE DICTIONARY OF THE ENGLISH LANGUAGE 132 (1966).

128. The Court then cites *Hazen Paper* in support of its "ordinary meaning" argument that "because of" means but for. *See Gross*, 129 S. Ct. at 2350 (citing Hazen Paper Co., v. Biggins, 507 U.S. 604, 610 (2003)). However, as noted above, *Hazen Paper* did not need to—and did not—decide which standard of causation to apply, given that there were no allegations of multiple factors. *See supra* text preceding note 189.

129. *See Gross*, 129 S. Ct. at 2350 (citing W. KEETON, ET AL., PROSSER AND KEETON ON THE LAW OF TORTS 265 (5th ed. 1984)).

130. *See* Katz, *supra* note 14, at 544, 549 (noting trends in modern tort law, including its use of a necessity-or-sufficiency standard instead of a simple necessity (but-for) standard).

131. *See* RESTATEMENT (SECOND) OF TORTS §432(2) (1965).

132. *See id.* (causation may be found in over-determined cases when defendant's conduct is independently sufficient to bring about the harm); *see also* RESTATEMENT (THIRD) OF TORTS § 27 (2005) (liability based on sufficiency in over-determined cases); Katz, *supra* note 14, at 521 n.122.

133. *See* Katz, *supra* note 14, at 544, 549.

134. *See* Katz, *supra* note 14, at 549.

the fact-finder might make the defendant pay only 50% of the plaintiff's damages.) In this way, neither party reaps a total windfall. Rather, windfall is allocated among the parties based on fault.

Arguably, this choice of causal standard—a necessity-or-sufficiency standard with damages allocation based on relative fault—was not available to the Court in *Gross*. After all, Congress nowhere mentions this standard in the legislative history of any of the disparate treatment statutes that use the "because of" formulation. And no court that I am aware of considers adopting such a standard. But my point here is merely that *Gross* adopted a normatively sub-optimal causal standard. If the goal of the law is to minimize windfall to one party, the simple necessity/but-for standard adopted by *Gross* is not as good as the necessity-or-sufficiency standard with damages apportionment used in modern tort law—the area of law *Gross* purports to use as a model.

Notably, modern employment anti-discrimination law provides a second-best alternative to the modern tort law standard, an alternative which still solves the windfall problem (albeit in a less equitable way than tort law's solution): The 1991 Act imposes liability with limited damages upon a showing of motivating factor causation when there is no but-for causation (that is, when the defendant prevails on the same action defense). This standard effectively divides damages in over-determined cases, thereby avoiding a complete windfall to either party. The 1991 Act standard does not apportion damages according to relative fault, like modern tort law does. But the 1991 Act standard arguably provides "rough justice."

Accordingly, Congress should, at the very least, extend the 1991 Act standard (motivating factor causation for liability, but-for causation for full damages) to the ADEA and other disparate treatment standards. Even better, Congress should adopt the modern tort law standard (motivating factor causation for liability, either sufficiency of necessity/but-for causation for damages, and apportionment of damages based on relative fault in over-determined cases) for all disparate treatment statutes.

CONCLUSION

The lesson to be gleaned from *Gross* seems to be two-fold. First, the Court seems completely uninterested in total unification; it intends to limit the benefits of the fairly good 1991 Act definition of "because of" to cases brought under the 1991 Act. Second, in cases outside of the 1991 Act, the Court seems determined to apply the worst possible

definition of "because of": but-for causation with no burden shifting.[135] This definition provides a windfall to defendants, fails to punish discriminators, under-deters discrimination, and places an undue burden of proof on plaintiffs. And there is every reason to believe that the Court will apply this flawed definition as broadly as possible, to all disparate treatment statutes other than the 1991 Act. It is therefore time for Congress to intervene.[136]

As this Article goes to press, Congress is in fact considering an amendment to the ADEA which would apply the 1991 Act definition of "because of" to ADEA cases.[137] The stated purpose of this proposed act is to overrule *Gross*. However, given the Court's apparent hostility to unification—particularly when it involves burden-shifting—the most important part of the proposed act may be a provision buried at the end of the Act, stating that the Act's causation requirement will apply to all federal anti-discrimination and anti-retaliation laws.[138] Without such a provision, even if Congress succeeds in amending the ADEA, the Court would likely apply the new standard only in ADEA cases, and continue to apply a *Gross*-like standard (but-for causation without burden-shifting) to other disparate treatment statutes.[139] Thus, in overruling *Gross*, Congress should apply the critical lesson of *Gross*: If Congress wants to apply a unified standard in all disparate treatment statutes, it needs to avoid amending one statute at a time. Instead, Congress should clearly express its intent to apply the new standard in all such statutes. In a post-*Gross* world, unification will need to come from Congress.

135. This is, admittedly, a slight overstatement. Certainly, requiring "sole cause," or "both necessity and sufficiency" would be normatively worse than the simple but-for standard adopted in *Gross*. However, neither of these two worse standards has ever been used in disparate law (and the first is likely precluded by the relatively clear legislative history of Title VII, see Price Waterhouse v. Hopkins, 490 U.S. 228, 241 n.7 (1989)). Among the true contenders, simple but-for is the worst.

136. *See* Corbett, *supra* note 5, at 9 (*Gross* "has made the need for [Congressional] intervention far more urgent. . . . *Gross* should make it abundantly clear that Congress must draft a clear blueprint addressing the issues rather than simply tearing down particular Supreme Court decisions.").

137. *See* Protecting Older Workers Against Discrimination Act, H.R. 3721, 111th Cong. (2009); Protecting Older Workers Against Discrimination Act, S. 1756, 111th Cong. (2009).

138. *See id.*, § 3(g)(5). It is worth noting that this provision also purports to apply the Act's causation requirements to constitutional anti-discrimination provisions. *Id.* It is far from clear that Congress has the power to set constitutional standards different from those determined by the Court.

139. A possible exception would be the ADA. *See supra* note 534.