BYU Law Review

Volume 49 | Issue 2 Article 1

Winter 12-27-2023

Full Issue

Brigham Young University Law Review

VOLUME 49 NUMBER 2

ARTICLES

NOTES & COMMENTS

Brigham Young University Law Review

VOLUME 49 NUMBER 2

BRIGHAM YOUNG UNIVERSITY LAW REVIEW (ISSN 0360-151X) will publish six issues in Volume 49 (December 2023 and January, February, March, May, July 2024). Business Office: J. Reuben Clark Law School, Accounting Office, 368 JRCB, Provo, Utah 84602-8000.

SUBSCRIPTIONS: The subscription rate for subscribers in the United States is $36 for one year, $65 for two years, and $90 for three years. Foreign subscriptions are $44 for one year, $81 for two years, and $114 for three years. All claims on issues not received must be made within three months of publication if within the United States, or within six months for foreign subscriptions. Subscriptions are entered for an entire year only and are payable in advance. All subscriptions are renewed automatically unless timely notice of cancellation is given. Please send all correspondence regarding subscriptions to Accounting Office, 368 JRCB, Provo, Utah 84602-8000. Issues older than one year may be obtained for $16.00 to $20.00 from William S. Hein & Co., Inc., 1285 Main Street, Buffalo, New York 14209-1987; phone (800) 828-7571. Back issues are also available at HeinOnline, http://heinonline.org.

SUBMISSIONS: The Law Review is pleased to accept electronic submissions for publication through Scholastica or by email: lawrev@lawgate.byu.edu. The Law Review no longer accepts hardcopy submissions for publication. All other communications pertaining to articles should be sent to the attention of the Managing Editor of Administration, 215 JRCB, Brigham Young University, Provo, Utah 84602-8000.

CONTENT: All citations conform to *The Bluebook: A Uniform System of Citation* (21st ed. 2020) and *The Chicago Manual of Style* (17th ed. 2017). Views expressed herein are to be attributed to their authors and not to the Brigham Young University Law Review or Brigham Young University, unless otherwise indicated.

Brigham Young University Law Review

VOLUME 49 NUMBER 2

LAW SCHOOL FACULTY

ANDERSON, CURTIS B. (2015), Teaching Professor of Law and Francis R. Kirkham Professor of Law. B.A., 1991, J.D., 1994, Brigham Young University.

ASAY, CLARK D. (2014), Hugh W. Colton Professor of Law. B.A., 2001, Brigham Young University; M.A., 2002, University of Cambridge; J.D., 2007, Stanford Law School.

AUGUSTINE-ADAMS, KIF (1995), Ivan Meitus Chair and Professor of Law. B.A., 1988, Brigham Young University; J.D., 1992, Harvard Law School.

BAUGHMAN, SHIMA BARADARAN (2023), Woodruff J. Deem Professor of Law. B.S., 2001, J.D., 2004, Brigham Young University.

BRAMBLE, CATHERINE E. (2023), Associate Professor of Law. B.A., 2002, J.D., 2005, Brigham Young University.

CLARK, ELIZABETH A. (2005), Associate Director for the International Center for Law and Religion Studies. B.A., 1994, J.D., 1997, Brigham Young University.

CLAYTON, WILLIAM W. JR. (2018), Associate Professor of Law and Francis R. Kirkham Professor of Law. B.A., 2006, Stanford University; M.B.A., 2010, Stanford Graduate School of Business; J.D., 2011, Yale Law School.

COLLINGS, JUSTIN M. (2013), University Academic Vice President and Professor of Law. B.A., 2006, Brigham Young University; J.D., 2010, Ph.D., 2013, Yale University.

DISHMAN, ELYSA (2016), Associate Professor of Law and Francis R. Kirkham Professor of Law. B.A., 2003, Brigham Young University; J.D., 2007, University of Virginia School of Law.

DOXEY, GARY B. (2005), Associate Director for the International Center for Law and Religion Studies. B.A., 1980, J.D., 1984, Brigham Young University; Ph.D., 1992, Cambridge University.

FEE, JOHN E. (2000), Professor of Law. B.A., 1992, Brigham Young University; J.D., 1995, University of Chicago.

FLEMING, J. CLIFTON JR. (1974), Ernest L. Wilkinson Chair and Professor of Law. B.S., 1964, Brigham Young University; J.D., 1967, George Washington University.

GALLI, CRAIG D. (2022), Distinguished Practitioner in Residence. B.A., 1982, M.A., 1984, Brigham Young University; J.D., 1987, Columbia University.

GEDICKS, FREDERICK M. (1990), Guy Anderson Chair and Professor of Law. B.A., 1977, Brigham Young University; J.D., 1980, University of Southern California.

GOTBERG, BROOK E. (2020), Professor of Law and Francis R. Kirkham Professor of Law. B.A., 2005, Brigham Young University; J.D., 2008, Harvard Law School.

GROW, LISA (2008), Howard W. Hunter Professor of Law. B.S., 1994, University of Utah; J.D., 1997, Harvard Law School.

HERNANDEZ, CARL III (2001), University Vice President of Belonging and Professor of Law. B.A., 1988, J.D./M.P.A., 1992, Brigham Young University.

JENNEJOHN, MATTHEW C. (2013), Professor of Law. B.A., 2002, Brigham Young University; MSc, 2004, London School of Economics; J.D., 2007, Columbia University School of Law.

JENSEN, ERIC TALBOT (2011), Robert W. Barker Professor of Law. B.A., 1989, Brigham Young University; J.D., 1994, University of Notre Dame; LL.M., 2001, The Judge Advocate General's School; LL.M., 2006, Yale Law School.

JONES, CREE (2020), Associate Professor of Law. B.S., 2006, Brigham Young University; M.A., 2009, John Hopkins University; J.D./M.P.P., 2013, University of Michigan; Ph.D., 2018, University of Chicago.

KYLE, KRISTIN B. GERDY (1995), Teaching Professor of Law and Director of Rex E. Lee Advocacy Program. B.A., 1992, J.D., 1995, Brigham Young University.

LEE, THOMAS R. (2022), Rex J. and Maureen E. Rawlinson Professor of Law. B.A., 1988, Brigham Young University; J.D., 1991, University of Chicago.

LEES, JESSICA (2022), Visiting Associate Professor. B.A., 2013, Brigham Young University; J.D., 2016, Boston University.

MOORE, DAVID H. (2008), Dean, Sterling and Eleanor Colton Endowed Chair in Law, and Professor of Law. B.A., 1992, J.D., 1996, Brigham Young University.

NIELSON, AARON L. (2013), Professor of Law. B.A., 2003, University of Pennsylvania; LL.M., 2006, University of Cambridge; J.D., 2007, Harvard Law School.

NÚÑEZ, D. CAROLINA (2007), Associate Dean for Research and Academic Affairs and Charles E. Jones Professor of Law. B.A., 2001, J.D., 2004, Brigham Young University.

PLAMONDON, STEPHANIE (2015), Professor of Law. B.S., 2001, University of Prince Edward Island; Ph.D., 2008, University of Utah; J.D., 2012, Harvard Law School.

REBEIRO, BRADLEY (2021), Associate Professor of Law. B.A., 2014, J.D., 2017, Brigham Young University; Ph.D., 2021, University of Notre Dame.

SCHARFFS, BRETT G. (1997), Rex E. Lee Chair, Director of International Center for Law and Religion Studies, and Professor of Law. B.S./B.A., 1986, M.A., 1987, Georgetown University; B.Phil., 1989, Oxford University; J.D., 1992, Yale Law School.

SHOBE, GLADRIEL D. (2016), Professor of Law. B.A., 2003, Brigham Young University; J.D., 2010, Yale Law School; LL.M., 2011, New York University.

SHOBE, JARROD E. (2016), Professor of Law. B.S./B.A., 2005, Arizona State University; J.D., 2011, Yale Law School.

SMITH, D. GORDON (2007), Ira A. Fulton Chair and Professor of Law. B.S., 1986, Brigham Young University; J.D., 1990, University of Chicago Law School.

STEELE, MICHALYN (2011), Associate Dean for Faculty and Curriculum and Marion G. Romney Professor of Law. B.A., 1992, M.A., 1994, Brigham Young University; J.D., 2001, Georgetown University Law Center.

THORLEY, DANE R. (2019), Associate Professor of Law. B.A., 2012, Brigham Young University; M.A., 2015, M.Phil., 2017, Columbia University; J.D., 2017, Yale Law School; Ph.D., 2019, Columbia University.

WILLIAMS, LUCY (2021), Associate Professor of Law. B.S., 2012, University of Utah; M.A., 2014, J.D., 2017, Ph.D., 2018, University of California, Los Angeles.

WORTHEN, KEVIN J (1987), University President and Hugh W. Colton Professor of Law. B.A., 1979, J.D., 1982, Brigham Young University.

ADJUNCT FACULTY

ARMSTRONG, RICHARD J.
BACALSKI, CHERISE
BARBER, W. BRAD
BARLOW, MATTHEW A.
BAUGHMAN, KRISTIN
BAZZELLE, SUCHADA
BEDNAR, STEVEN C.
BOONE, ANNIKA
CHRISTENSEN, MARIAH J.
COOK, MIKAELA
CROWTHER, COLE
CRUMP, CHRISTOPHER
DISHAMN, LAKE
DUTCHER, DAVID
EVANSON, BLAINE
EVERETT-GARCIA, JESSICA
FLOWERDAY, MARISSA
FOWLES, JOHN
GALLACHER, LESLIE
GLADE, LEEANN
GOODWIN, SCOTT D.
GRAHAM, ATHELIA
GRIFFITH, SUSAN W.
HALL, MANAL
HIGBEE, JOSH
HINCKLEY, JOSEPH
HORTON, KENNETH
JOHNSON, FREYJA
KELSON, STEPHEN
LORIMER, BRENT
MATIS, GREGORY J.
MCBRIDE, RYAN
MCCLURG, JOHN
MCCULLOUGH, LEE S.
MERRILL, ADAM
MOESEL, MIKE
MOTT, LILY
NELSON, RYAN D.
NEILSEN, JOHN
NOEL, SAMUEL
NUTTALL, TRACY
OWEN, STANFORD B.
PEAD, HON. DUSTIN B.
PEARSON, JOI
PERRY, AUDREY
PFROMMER, SARA
PHILLIPS, SPENCER D.
POMEROY, ADAM
RANE, KATIE
SCHAERR, GENE C.
THOMAS, SUSANNAH
THOMPSON, DOUGLAS
TIPPETT, ROBERT S. JR.
TODD, DAVID R.
VOGELER, ERIC
WAHLQUIST, ADAM
WALKER, DANNY
WALKER, JEFFREY N.
WEILER, TODD D.
WHITTAKER, JUSTIN
WILLIAMS, JAMIE
WISE, JANE
YOUNKER, NORM

Brigham Young University Law Review

VOLUME 49 NUMBER 2

Associate Editors

TAYLIN ANTONICK
JESSICA ASHBY
KEVIN AUMAN
KESSLER BAKER
HANNAH BIESINGER
LINDI BRAY
MORGAN BRONSON
LIESA BRUIN
JOSHUA BUCKNER
SARAH CALVERT
RYAN CHENEY
ABRAM DORROUGH
ANDREW FELLOWS
NATASHA FLORA
DAVID GRIFFITHS
ELIZA GUTIERREZ
JACOB HANSEN
LANDON HARDY
CLAIRE HAWKINS
NAOMI HILTON
ALYSSA HUNZEKER
TAYLOR JEANFREAU
BENSON KILLPACK
REBECCA KNOWLES
MALCOLM LAMB
REBEKAH LEAVITT-HATCH
NICK LOTT
CASSIDY MCCLEARY
JOHN MCDONALD
COLTER MILLER
JONATHAN MIX
KEIL NICHOLAS
BRIDGER PAGE
TAYLOR PERCIVAL
RYAN QUADE
BRADEN REESE
MARSHALL RINGWOOD
ELIZA SMITH-DRIGGS
SONJA SMITH SHINEDLING
BONNIE STEWART
RIVER TANNER
WESLEY WHITE

Hidden Contracts

Shmuel I. Becher & Uri Benoliel†*

Transparency is a promising means for enhancing democratic values, countering corruption, and reducing power abuse. Nonetheless, the potential of transparency in the domain of consumer contracts is untapped. This Article suggests utilizing the power of transparency to increase consumer access to justice, better distribute technological gains between businesses and consumers, and deter sellers from breaching their consumer contracts while exploiting consumers' inferior position.

In doing so, this Article focuses on what we dub "Hidden Contracts." Part I conceptualizes the idea of hidden contracts. It first defines hidden contracts as consumer form contracts that firms unilaterally modify and subsequently remove from the public sphere, despite being binding on consumers. Thereafter, the Article delineates the considerable social costs of hidden contracts.

Given these social costs, Part II discusses our empirical study of hidden contracts. The results of this study indicate that leading firms that supply goods and services to billions of online consumers worldwide routinely employ hidden contracts to the detriment of consumers and society. Against this background, Part III proposes introducing a novel contract transparency duty. It further explains how to design this duty to counter firms' incentive to employ hidden contracts. Next, Part IV tackles key objections to our proposal. Concluding remarks follow.

* Professor of Law, Victoria University of Wellington; Lee Kong Chian Visiting Professor of Law, Yong Pung How School of Law, Singapore Management University.

† Professor of Law, College of Law and Business. J.S.D. (University of California, Berkeley); LLM (Columbia University). We thank Yonathan Arbel, Oren Bar-Gill, Sinai Deutch, Chris Drahozal, Grace Henry, Dave Hoffman, David Horton, Mark Budnitz, Roy Shapira, Amy Schmitz and Jeff Sovern for their comments on earlier drafts, Efrat Avivi for able research assistance, and Victoria University of Wellington and the College of Law and Business for financial support.

CONTENTS

INTRODUCTION

Consumer standard form contracts—the most pervasive type of contracts—govern much of our everyday lives.[1] One enters into a standard form contract when using social media, signing up for an

1. *See, e.g.*, W. David Slawson, *Standard Form Contracts and Democratic Control of Lawmaking Power*, 84 HARV. L. REV. 529, 529 (1971) ("Standard form contracts probably account for more than ninety-nine percent of all of the contracts now made."); *see also* Robert A. Hillman & Jeffery J. Rachlinski, *Standard-Form Contracting in the Electronic Age*, 77 N.Y.U. L. REV. 429, 435 (2002) (noting the omnipresence of consumer form contracts, their wide usage in e-commerce, and some of the relationships they govern).

email account or a ridesharing service, booking a flight or a hotel, renting a car, opening a bank account, purchasing insurance, joining a gym, connecting to public Wi-Fi, or buying furniture at a local store. Consumer form contracts are a constant presence, and Terms and Conditions (or Ts & Cs, Terms of Service, Terms of Use (ToU), or simply Terms) govern the behavior, rights, and obligations of billions of individuals worldwide.[2]

Ubiquity notwithstanding, consumers often think nothing of these contracts until a dispute arises and they discover that the odds are weighed heavily against them.[3] Indeed, academics have been sounding the alarm about various problematic aspects of consumer standard form contracts for decades.[4] Scholars question the validity of consumers' assent to standard form contracts, noting the degradation of consumer rights, lack of choice, and consumers' inferior bargaining power.[5] Some opine that consumer form contracts grant firms an (absolute or excessive) ability to one-sidedly draft, design, amend, and resolve disputes pertaining to

2. For example, the mega social networks Facebook, YouTube, Instagram, and TikTok have approximately 3, 2.5, 2, and 1.1 billion active users, respectively. *See Most Popular Social Networks Worldwide as of January 2023, Ranked by Number of Monthly Active Users*, STATISTA, https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/ (last visited Oct. 2, 2023). Likewise, the popular online retailers Amazon, eBay, and Walmart have approximately 3.2 billion, 589 million, and 581 million monthly visits, respectively. *See Most Visited Online Retail Websites Worldwide in 2022, by Average Monthly Traffic*, STATISTA, https://www.statista.com/statistics/274708/online-retail-and-auction-ranked-by-worldwide-audiences (last visited Oct. 2, 2023).

3. For one illustrative example, see Katie Benner, *Federal Judge Blocks Racial Discrimination Suit Against Airbnb*, N.Y. TIMES, Nov. 1, 2016, at B5 ("[A] federal judge ruled that the company's arbitration policy prohibited its users from suing.").

4. For two seminal early examples, see Friedrich Kessler, *Contracts of Adhesion – Some Thoughts About Freedom of Contract*, 43 COLUM. L. REV. 629 (1943) (explaining that standard form contracts are contracts of adhesion, thus deviating from fundamental assumptions of traditional contract law); KARL N. LLEWELLYN, THE COMMON LAW TRADITION: DECIDING APPEALS 362–71 (1960) (analogizing standard form contracts to laying one's head into a lion's mouth).

5. The literature here is extensive. For a few examples, see Lewis A. Kornhauser, Comment, *Unconscionability in Standard Forms*, 64 CALIF. L. REV. 1151, 1162 (1976) ("Most clauses of standard form contracts are candidates for nonenforcement."); Arthur Allen Leff, *Contract as Thing*, 19 AM. U. L. REV. 131, 143 (1970) (submitting that contracts of adhesion are drafted by one party only and are not a result of a cooperative negotiation); MARGARET JANE RADIN, BOILERPLATE: THE FINE PRINT, VANISHING RIGHTS, AND THE RULE OF LAW (2013) (detailing the urgent need to improve oversight of boilerplate contract terms).

consumer contracts.[6] Others warn that consumers do not, and cannot, read complex and lengthy standard form contracts.[7] Vast literature wrestles with this long-lasting "no-reading problem" and the asymmetric information it facilitates.[8]

In the domain of Business-to-Consumer (B2C) relationships and elsewhere, one of the principles that can counter and disincentivize exploitative behavior is transparency.[9] Transparency plays an important role in many legal and nonlegal domains.[10] It enhances

6. *See, e.g.*, David Horton, *The Shadow Terms: Contract Procedure and Unilateral Amendments*, 57 UCLA L. REV. 605, 667 (2010) (discussing firms' ability to "unilaterally amend their dispute resolution clauses again and again"); Kessler, *supra* note 4; Slawson, *supra* note 1; Eyal Zamir, *Commonsense Consent and Contract Law*, JOTWELL (Dec. 5, 2022) (reviewing Joanna Demaree-Cotton & Roseanna Sommers, *Autonomy and the Folk Concept of Valid Consent*, 224 COGNITION 105065 (2022)), https://contracts.jotwell.com/commonsense-consent-and-contract-law (noting the weak regulation of consumer form contracts in the United States).

7. *See, e.g.*, Yannis Bakos, Florencia Marotta-Wurgler & David R. Trossen, *Does Anyone Read the Fine Print? Consumer Attention to Standard-Form Contracts*, 43 J. LEGAL STUD. 1 (2014) (establishing empirically that virtually no consumers read End User License Agreements); Melvin A. Eisenberg, *The Limits of Cognition and the Limits of Contract*, 47 STAN. L. REV. 211, 243 (1995) (explaining that consumers who typically face one-shot transactions will not accord much attention to standardized terms); Melvin A. Eisenberg, Comment, *Text Anxiety*, 59 S. CAL. L. REV. 305 (1986) (opining that information overload and language complexity will deter consumers from reading form contracts); Todd D. Rakoff, *Contracts of Adhesion: An Essay in Reconstruction*, 96 HARV. L. REV. 1173, 1179 (1983) (highlighting the consensus among academics that the adhering party is unlikely to read the standard terms before accepting them). We will return this important issue *infra* Section IV.D.

8. *See, e.g.*, Ian Ayres & Alan Schwartz, *The No-Reading Problem in Consumer Contract Law*, 66 STAN. L. REV. 545 (2014) (proposing that firms engage in "term substantiation" to learn whether consumers hold accurate beliefs about their agreements); Shmuel I. Becher, *Asymmetric Information in Consumer Contracts: The Challenge That Is Yet to Be Met*, 45 AM. BUS. L.J. 723, 773 (2008) (explaining how the no-reading problem results in contractual information asymmetries that market forces cannot correct); Amy J. Schmitz, *Pizza-Box Contracts: True Tales of Consumer Contracting Culture,* 45 WAKE FOREST L. REV. 863 (2010) (exploring the problem of rolling contracts, where consumers get the terms of their contracts only post purchase).

9. *See, e.g.*, Shmuel I. Becher & Uri Benoliel, *Dark Contracts,* 64 B.C. L. REV. 55 (2023) (proposing a transparency framework to scrutinize and police firms' contractual behavior).

10. For a detailed discussion, see ARCHON FUNG, MARY GRAHAM & DAVID WEIL, FULL DISCLOSURE: THE PERILS AND PROMISE OF TRANSPARENCY (2007). *See also* Stephen Kosack & Archon Fung, *Does Transparency Improve Governance?*, 17 ANN. REV. POL. SCI. 65, 68 (2014) (recognizing various ways to employ transparency); David Weil, Mary Graham & Archon Fung, *Targeting Transparency*, 340 SCI. 1410 (2013) (explaining, among other things, how policymakers can use transparency as a tool to empower third parties to provide market participants with valuable and accessible information).

democratic values, counters corruption, and reduces power abuse.[11] Transparency also facilitates equality, fosters trust, and promotes fairness.[12] Few will doubt the positive attributes and the potential of transparency to improve lives, markets, and institutions.

Unsurprisingly, therefore, giant websites—including the most famous and successful social media platforms—often state that they are committed to the principle of transparency. To illustrate, Facebook declares on its website that it is "committed to transparency, control and accountability."[13] Similarly, Twitter states that it "was founded on a commitment to transparency."[14] Along these same lines, TikTok asserts, "We work to earn and maintain trust through ongoing transparency into the actions we take to safeguard our platform . . . ."[15]

Firms' commitment to transparency is socially important. It builds trust between these websites and their users.[16] It allows businesses to establish a reputation for openness with the public.[17]

11. *See Our Story*, TRANSPARENCY INT'L, https://www.transparency.org/en/our-story (last visited Oct. 2, 2023).

12. *See, e.g.*, Perrine Toledano, *Paper on the Business Case for Transparency*, COLUM. CTR. ON SUSTAINABLE INV. STAFF PUBL'NS, June 2012, at 1, 9 (associating transparency with accountability, fairness, and sustainability); Archon Fung, Mary Graham & David Weil, *The Political Economy of Transparency: What Makes Disclosure Policies Sustainable?*, ASH CTR. FOR DEMOCRATIC GOVERNANCE & INNOVATION (2002), https://ash.harvard.edu/publications/political-economy-transparency-what-makes-disclosure-policies-sustainable.

13. *See, e.g.*, *Facebook's Commitment to Data Protection and Privacy in Compliance with the GDPR*, META (Jan. 29, 2018), https://bit.ly/2EGyMvB.

14. *Twitter Russia Transparency Report*, TWITTER (Feb. 2022), https://bit.ly/3Cg0wmW.

15. Vanessa Pappas, *Strengthening Our Commitment to Transparency*, TIKTOK (Jul. 27, 2022), https://bit.ly/3UODBGI. For additional examples, see *Discount Transparency Initiative*, MICROSOFT, https://bit.ly/3ClBhQk (last visited Oct. 2, 2023) ("Microsoft is dedicated to conducting business responsibly around the world and to meeting the highest legal and ethical standards. Improved transparency is an essential component of these efforts."); *Transparency Report*, PINTEREST, https://policy.pinterest.com/en/transparency-report (last visited Oct. 2, 2023) ("We're committed to providing greater transparency into how we keep Pinterest safe and positive . . . .").

16. *See TikTok's New Transparency Reports and Transparency Center*, TIKTOK (DEC. 2, 2021), https://bit.ly/3dZeRdY ("To build trust through transparency, we began releasing reports in 2019 . . . ."); *Transparency Report*, ZOOM (Jul. 7, 2023), https://explore.zoom.us/docs/en-us/trust/transparency.html ("At Zoom, transparency is critical to building trust . . . .").

17. *See Transparency Reporting*, TSPA, https://www.tspa.org/curriculum/ts-fundamentals/transparency-report (last visited Oct. 2, 2023) ("Transparency reporting can be important for building trust and establishing a reputation for openness with the public.").

It helps companies to foster accountability in their work.[18] It may also reduce policymakers' vigilance towards firms and shield firms from public and legal scrutiny. If sellers voluntarily adopt transparency as a guiding principle, legal intervention to promote transparency and protect the public becomes less justified.

Given firms' stated commitment to the transparency principle, a fundamental yet imperative question arises: Are websites systematically transparent about the content of their legal rules? Put simply, do consumers have access to the form contracts they accepted when entering into their relationships with the firm? Or do firms opt for non-transparency and hide previous—yet relevant—versions of their contracts from consumers?

Addressing these questions, this Article empirically investigates the contractual practices of 100 highly popular mega sites, such as Google, Facebook, Instagram, TikTok, Twitter, and Amazon. In particular, it examines whether highly popular sites use nontransparent practices that characterize what we dub "hidden contracts." Hidden contracts, as this Article defines them, are mass, standard form contracts whose tracks are blurred to prevent consumers from accessing them. In other words, hidden contracts are agreements consumers cannot find on the online vendor's website once a conflict or dispute arises—typically when they need them most. Consequently, hidden contracts go beyond the concerns that consumer contracts are hard to understand or may have changed; rather, they imply that the terms are gone or inaccessible.

To illustrate the issue, consider the case of the Palmers from Utah. Mr. Palmer ordered a desk ornament and keychain from an online retailer, KlearGear.[19] When the retailer did not deliver, the

18. Chris Sonderby, *Transparency Report, Second Half 2021*, META (May 17, 2022), https://about.fb.com/news/2022/05/transparency-report-h2-2021 ("That's why we publish biannual transparency reports to provide detail on the numbers and maintain accountability in our work.").

19. Many outlets reported on the Palmers' story and subsequent lawsuit against KlearGear. *See, e.g.*, *Palmer v. KlearGear.com*, PUB. CITIZEN, https://www.citizen.org/litigation/palmer-v-kleargear-com (last visited Oct. 2, 2023); Steven Nelson, *Retailer That Fined Couple $3,500 for Negative Review Hit with Lawsui*t, USNEWS (Dec. 18, 2013, 7:00 am), https://www.usnews.com/news/articles/2013/12/18/retailer-that-fined-couple-3500-for-negative-review-hit-with-lawsuit; Olivia Sorrel-Dejerine, *A Case for Reading the Small Print,* BBC (Nov. 18, 2013), https://www.bbc.com/news/blogs-magazine-monitor-24992518; Jacob Goldstein & Alexi Horowitz-Ghazi, *Terms of Service*, NPR (Mar. 4, 2020, 7:11 PM),

Palmers attempted to find out what went wrong.[20] Upon realizing she could not reach the business by phone, Mrs. Palmer posted an online review about the business, sharing her negative experience.[21]

More than three years later, the online retailer, KlearGear, contacted the Palmers, demanding the removal of the negative posting. It also threatened the couple with a $3,500 fine.[22] The message from KlearGear cited the firm's terms and conditions, pointing to a non-disparagement clause, which prohibits consumers from publishing negative feedback about the company.[23]

One of the key issues in this case was whether this term was part of the original contract between the parties.[24] On the one hand, KlearGear stated that the clause had been present in 2008 when the Palmers entered the contract. On the other hand, the Palmers argued that according to Internet Archive, an American digital library, the clause had not been part of the original contract and was added to the site in June 2012.[25] In other words, KlearGear inserted this term *after* the Palmers entered into the contract.[26]

When the Palmers refused to pay the fine or remove the negative review, KlearGear reported them to several credit

https://www.npr.org/2020/03/04/812264543/episode-976-terms-of-service. The story evolved to a lawsuit, with its own Wikipedia entry. *See Palmer v. Kleargear.com,* WIKIPEDIA, https://en.wikipedia.org/wiki/Palmer_v._Kleargear.com (last visited Oct. 2, 2023).

20. *See* PUB. CITIZEN, *supra* note 19 ("The gifts never arrived, and [Mr. Palmer's] attempts to contact KlearGear.com were unsuccessful.").

21. *See id.* (noting that Mrs. Palmer "posted a negative review on RipoffReport.com").

22. *See id.*; Cyrus Farivar, *KlearGear Must Pay $306,750 to Couple That Left Negative Review*, ARS TECHNICA (June 25, 2014, 8:10 PM), https://arstechnica.com/tech-policy/2014/06/kleargear-must-pay-306750-to-couple-that-left-negative-review (describing KlearGear's initial demand to the Palmers to remove their review or pay the large fine).

23. Farivar, *supra* note 22 ("[Mr.] Palmer received an e-mail demanding that the review be deleted within seventy-two hours or that he pay $3,500, as he was in violation of the company's 'non-disparagement clause' of its terms of service.").

24. Farivar, *supra* note 22 (noting that the clause "did not appear in the Terms of Sale and Use that the Palmers had agreed to when they placed their order in 2008").

25. Complaint at clause 19, Palmer v. Kleargear.com, No. 13-cv-00175 (D. Utah Dec. 18, 2013), ("Based on past versions of KlearGear's Terms of Sale and Use available at the Internet Archive, https://archive.org, it appears that KlearGear's Terms of Sale and Use did not include a 'non-disparagement clause' until sometime after April 28, 2012."). *See also* Tim Cushing, *Online Retailer Says if You Give It a Negative Review It Can Fine You $3,500,* TECHDIRT (Nov. 14, 2013, 8:58 AM), https://www.techdirt.com/2013/11/14/online-retailer-slaps-unhappy-customers-with-3500-fee-violating-non-disparagement-clause.

26. Complaint, *supra* note 25.

bureaus.[27] As a result, the Palmers were denied credit, had loans delayed, and could not enlist the necessary funds to fix their broken furnace.[28] The Palmers sued KlearGear,[29] and the court entered a default judgment in favor of the Palmers.[30] Later, Congress banned the use of "non-disparagement clauses" (known more generally as "gag clauses") by statute, introducing the Consumer Review Fairness Act.[31]

Nonetheless, consumers' attempts to exercise their rights do not often end as fortuitously as in the Palmers' case. Most consumers find it challenging to confront firms that unilaterally change the contract terms and obfuscate the original terms of the agreement. Firms typically incorporate contractual change-of-terms mechanisms that allow them to unilaterally modify the original contract at will, for any reason, and at any time.[32] This practice can exacerbate consumers' difficulties. As a matter of fact, online forums provide multiple anecdotal examples of customers and users struggling, with no success, to find the relevant terms and conditions.[33]

27. *See* PUB. CITIZEN, *supra* note 19 ("When the Palmers refused to pay . . . KlearGear.com reported the supposed 'debt' to the credit reporting agencies. More than a year later . . . this 'debt' still mars John Palmer's credit.").

28. *Id.* ("[T]he Palmers have been turned down for credit, and had their car loan delayed and paid a higher interest rate on it . . . ."); Eugene Volokh, *$300,000 Damages Award Against KlearGear, the Company that Billed Customers for $3,500 Because They Posted a Negative Review,* WASH. POST: THE VOLOKH CONSPIRACY (June 26, 2014, 12:48 PM) (stating that the Palmers "spent weeks without heat in their home . . . when their furnace broke and they were unable to obtain a loan to replace it").

29. Complaint, *supra* note 25.

30. Palmer v. Kleargear.com, 13-cv-00175 (D. Utah filed Dec. 18, 2013). *See also* PUB. CITIZEN, *supra* note 19 (describing the default judgment and damages award).

31. Consumer Review Fairness Act of 2016, Pub. L. No. 114-258, 130 Stat. 1355.

32. *See, e.g.*, Shmuel I. Becher & Uri Benoliel, *Sneak in Contracts*, 55 GA. L. REV. 657 (2021) (finding that the vast majority of popular online website incorporate a non-transparent change-of-terms clause in their consumer form contracts).

33. *See, e.g.*, Live Better, *We Offer A $900 Reward to Anyone Who Finds a 2017 Version of Epidemic Sound Terms of Service,* LIVEBETTERMEDIA (Mar. 31, 2020), https://tinyurl.com/38pc5rvh ("Epidemic Sound is hiding the previous versions of its Terms of Service . . . . [A]nd if you ask them to re-release these past versions of the documents that once governed Epidemic Sound's website, services and copyrights, they ignore your requests, even if the requests are made by a lawyer."); tjayhawk3231, Comment to *About Skins Being a 'Rental'*, REDDIT (2022), https://www.reddit.com/r/PlayAvengers/comments/soxr29/about_skins_being_a_rental/?rdt=56272 ("I can't find the original terms for Marvel Heroes."); Maxine H, *Recent Comments & Queries from Other Timeshare Owners*, TIMESHARE ADVICE NETWORK (Aug.

As the Palmers' case illustrates and this Article elaborates below,[34] it is rather beneficial for consumers—as well as for consumer organizations, watchdog groups, policymakers, enforcement agencies, the media, and adjudicators—to easily know what a consumer contract, including its previous and original versions, says. But despite the fundamental nature of this issue, we are unaware of any studies exploring it. This Article marks the first attempt to systematically examine whether consumers can locate the original contract's terms and conditions that govern their relationships with suppliers.

The key contribution of this Article is threefold. First, it empirically addresses an important question that the literature has neglected: Whether the content of online consumer contracts is available and transparent. Second, this Article joins the call to channel further scholarly and regulatory attention to the hurdles consumers face ex post, when attempting to insist on their rights or confront firms.[35] Third, this Article connects three premises—access to justice, firms' non-transparent behaviors, and the potential and perils of technology—to make a novel argument regarding the regulation of popular online websites, platforms, and services.[36] It argues that big online firms often employ non-transparent tools

18, 2017), https://www.timeshareadvicenetwork.co.uk/your-stories ("I can't find the original contract."); Doingmybest, Comment to *FirstPlus to Elderbridge*, CONSUMER ACTION GROUP (Apr. 25, 2018), https://www.consumeractiongroup.co.uk/topic/410000-firstplus-to-elderbridge ("I can't find the original contract and terms.").

34. *See* discussion *infra* Section I.B.

35. This attention, in turn, highlights the discriminatory means firms employ when dealing with consumers. *See, e.g.*, Yonathan A. Arbel & Roy Shapira, *Theory of the Nudnik: The Future of Consumer Activism and What We Can Do to Stop It*, 73 VAND. L. REV. 929 (2020) (documenting how assertive and vocal consumers may discipline firms and evaluating those customers' roles); Shmuel I. Becher & Tal Z. Zarsky, *Minding the Gap*, 51 CONN. L. REV. 69 (2019) (discussing the motivations and implications of firms that draft strict form contracts ex ante yet display leniency toward consumers ex post); Amy J. Schmitz, *Access to Consumer Remedies in the Squeaky Wheel System*, 39 PEPP. L. REV. 279 (2012) (explaining how firms discriminate among consumers, allowing only relatively few vocal and informed consumers to access remedies). *See also* Manisha Padi, *Contractual Inequality,* 120 MICH. L. REV. 825 (2022) (exploring the magnitude of ex post contractual inequality in the context of residential mortgage contracts); Rory Van Loo, *The Corporation as Courthouse*, 33 YALE J. ON REG. 547 (2016) (highlighting the role of internal corporate mechanisms in responding to, resolving, and mitigating consumer complaints).

36. *See, e.g.*, John M. Newman, *Antitrust in Digital Markets*, 72 VAND. L. REV. 1497 (2019) (arguing that digital markets merit a uniquely interventionist approach); Roy Shapira, *The Challenge of Holding Big Business Accountable,* 44 CARDOZO L. REV. 203 (2022) (assessing how size and market power create governance problems and how the law can respond).

while eroding consumer access to justice. In doing so, this Article sheds much-needed light on the untapped potential of transparency in the domain of consumer contracts.

The structure of this Article is as follows. Part I provides the theoretical context for the empirical test of this study. It defines hidden contracts and examines their social costs. Part II presents the empirical test of this study. It reviews the data that underlines the empirical examination and discusses its methodology. It then details the results of this study, which indicate that highly popular websites too often apply practices that yield hidden contracts. Part III discusses normative policy and legal implications. It suggests imposing a transparency duty on firms and considers private and public enforcement measures to tackle the challenge of hidden contracts. Part IV discusses potential criticism. Concluding remarks follow.

I. THEORETICAL BACKGROUND

This Part conceptualizes the idea of hidden contracts. Section A defines the two components of such contracts, illustrating them by reference to Amazon. Thereafter, section B delineates the social costs of hidden contracts.

A. The Concept of Hidden Contracts

This Article defines hidden contracts as standard form agreements that oblige consumers despite being concealed from them. More specifically, a supplier that utilizes hidden contracts applies two practices: First, it does not provide a copy of the agreement to consumers after they agree to it. Second, the supplier removes the previous version of the contract from its public sphere after unilaterally amending it.[37] We use the term "hidden" to emphasize that whereas firms have records of these previous contract versions in their databases, consumers cannot find or access them.

To illustrate, take the typical consumer experience with Amazon. When consumers sign up for an Amazon account, they

37. Importantly, consumer contracts are often amended by suppliers. For example, Twitter's contract terms were amended at least fifteen times from the year 2009 to 2022. *See Previous Terms of Service*, TWITTER, https://twitter.com/en/tos/previous (last visited Sep. 15, 2023).

are required to provide their contact information, such as their email address.[38] Following that, consumers must agree to Amazon's contractual terms of use ("original version").[39] However, after consumers agree to the contract, Amazon does not send a copy of the original version to consumers.[40] Occasionally, Amazon will unilaterally amend its previous contractual terms of use,[41] creating a modified version of its form contract.[42] However, when it makes such a unilateral amendment, it fails to publish or link to the original version on its terms of use webpage.[43] As a result, typical consumers, who accepted the original version, cannot know what it says and are unable to find it.

B. The Social Costs of Hidden Contracts

Hidden agreements are undesirable. In essence, hidden contracts increase the risks that consumers will not know their rights and will avoid action. As a result, businesses will be under-deterred and inefficiently breach their standard form contracts at the expense of consumers.

This section clarifies this argument in three steps: First, it explains why a supplier may breach its contract inefficiently. Second, it discusses a major social mechanism against inefficient breaches by suppliers: consumers' activism and their ability to sue breaching suppliers. Finally, it submits that hidden contracts erode consumer activism; i.e., hidden contracts deter consumers from confronting firms, complaining about their behavior, and filing lawsuits against suppliers who breach their contracts. This reality, in turn, entails that hidden contracts assist firms in breaching their contracts inefficiently. Hidden contracts, therefore, harm consumers

38. *Create Account*, AMAZON, https://amzn.to/3RJYAZ0 (last visited Oct. 3, 2023).

39. *Id.* (Amazon states following its sign-up button: "By creating an account, you agree to Amazon's Conditions of Use . . . .").

40. To confirm this reality, one of the authors signed up with Amazon in September 2022. During the sign-up process the author provided his email address and accepted Amazon's contract terms. However, Amazon did not email (or otherwise send) the author a copy of the contract he agreed to.

41. *Conditions of Use*, AMAZON, https://amzn.to/3QJ32Gm (last updated Sep. 14, 2022) (According to Amazon, it last amended its terms on September 14, 2022.).

42. Of course, many companies amend their contracts more than once, hence creating multiple versions of the consumer form contract they offer throughout time.

43. *Supra* note 41.

and undermine the ability of market forces and information flows to discipline sellers and benefit society.

C. Suppliers May Breach Their Contracts Inefficiently

Firms, as profit-maximizers, have a basic incentive to save costs when possible.[44] Accordingly, when firms predict that the costs of executing their standard form contracts exceed the costs of breaching them, they are likely to breach the agreement.[45] That is, firms may focus only on their own efficiency curve and ignore consumers' interests.

An example may clarify. Assume that a social network bans a user from the network based on its assumption that the user breached its rules of conduct.[46] Furthermore, assume that according to the network's contract, the network must provide any banned consumer with a clear and detailed explanation about why they were banned.[47] In such a scenario, if the social network predicts its costs of explanation are higher than the costs of failing to explain, it is likely to choose the latter. Therefore, the network will ban the user without providing any reasoning whenever the cost of providing an explanation surpasses the expected harm to the network.[48] By banning the user without giving concrete justifications, the network may save—from its self-interested

44. Jae Sung Lee, *Towards a Development-Oriented Multilateral Framework on Competition Policy*, 7 SAN DIEGO INT'L L.J. 293, 303 (2006) ("Competition among firms sharpens incentives to cut cost."); Wulf A. Kaal, *Blockchain Technology for Good*, 17 U. ST. THOMAS L.J. 878, 883 (2022) ("[C]apitalism suggests that market incentives encourage firms to cut costs.").

45. Thomas S. Ulen, *Happiness, Technology, and the Changing Employment Relationship*, 19 EMP. RTS. & EMP. POL'Y J. 61, 66 (2015) ("[R]ational breachers are very likely to breach only when it is more efficient to breach than to perform."); *cf.* Jason N.E. Varuhas, *One Person Can Make A Difference: An Individual Petition System for International Environmental Law*, 3 N.Z.J. PUB. & INT'L L. 329, 335 (2005) ("[S]tates are less likely to breach their obligations if the costs of non-compliance are high.").

46. *See, e.g.*, King v. Facebook, Inc., 572 F. Supp. 3d 776, 781 (N.D. Cal. 2021) (Facebook banned a user since it believed that the user "did not follow [Facebook] Community Standards.").

47. *Cf. id.* at 789 (Court rules, based on the contract text, that the implied covenant of good faith obliges Facebook "to provide [to the user] at least some information in addition to the fact that the account has been suspended or terminated—e.g., enough information about why the account was suspended or terminated.").

48. *Id.* at 790 ("Based on the Court's analysis above, Ms. King [a Facebook user] has a claim for breach of contract (or breach of the implied covenant) based on Facebook's disabling of her account, as well as the failure to provide a more specific explanation as to why the account was disabled.").

perspective—the costs of explanation; that is, the costs of articulating the reasons for the ban.

However, while a breach of contract (i.e., not providing the explanation) may be beneficial for the firm, it may be inefficient from a broader social perspective. That is, the total social costs of the breach may outweigh the supplier's costs of providing the explanation. Going back to the example of the social network, when the network breaches the agreement with its banned consumer by failing to explain the reasons for the customer's termination, the overall social costs may be significant.[49]

For starters, the probability of erroneous contract-banning of an innocent consumer who did not breach the network's rules increases.[50] In other words, lack of reasoning increases the risk that the banning process applied by the supplier will be hasty and not founded on accurate, adequately investigated facts and sound principles of law.[51] Furthermore, a lack of reasoning may also impair the consumer's capacity to easily and fully understand the firm's banning decision and effectively contest it if it is wrong.[52]

In addition, the firm's erroneous banning of the consumer may generate nontrivial costs to the consumer, which the firm is unlikely to internalize.[53] Notably, these costs include the loss of the consumer's prior investment (e.g., time and resources) in trying to create and maintain a relationship with other network users.[54] Moreover, the banned user may experience embarrassment, mental anguish, and emotional distress due to an unjust banning by the firm.[55] On top of that, the consumer may incur switching costs when pursuing a similar service or product elsewhere.[56]

Finally, the firm's motivation to ban a consumer may (at times, erroneously) be based on discriminatory and non-transparent

49. Uri Benoliel & Shmuel I. Becher, *Termination Without Explanation Contracts*, 2022 U. ILL. L. REV. 1059, 1070–84 (2022).

50. *Id.* at 1075–78.

51. *Id.* at 1075.

52. *Id.* at 1077. *See also* King v. Facebook, Inc., *supra* note 46, at 789 ("[I]t is plausible that Facebook is obligated to provide at least some information in addition to the fact that the account has been suspended or terminated—e.g., enough information about why the account was suspended or terminated such that an 'appeal' could properly be made . . . .").

53. Benoliel & Becher, *supra* note 9, at 1078–84.

54. *Id.* at 1078–79.

55. *Id*. at 1081–83.

56. *Id*. at 1083–84.

factors that undercut societal values such as inclusion, equality, and voice.[57] Thus, such banning may disproportionately harm vulnerable consumers while eroding imperative societal values.[58] Nevertheless, since the firm does not fully internalize these costs, it may have a profit incentive to ban consumers without providing an explanation, despite this being a breach of contracts.

Against this background, consumers' ability to discipline firms and sue wrongdoers is of paramount importance. We turn to that next, examining the interplay between hidden contracts and consumers' propensity to complain, air their grievances, and litigate their cases.

1. Consumer Activism as a Mechanism Against Inefficient Breach

While suppliers may have a basic incentive to breach their contracts at the expense of consumers, society has an important mechanism that may reduce the occurrence of such inefficient breaches. Particularly, society allows consumers to file a lawsuit against a supplier that breached its contract. Such lawsuits may deter suppliers, ex ante, from breaching the contract.[59] This deterrence effect is due to the high monetary and reputational costs that consumer lawsuits may cause to a breaching supplier. By and large, a similar analysis applies to consumer activism that takes other shapes: complaining to the mass media and consumer organizations, sharing experiences on social media platforms, or posting online reviews. Though all these tools are not magic bullets,

57. *Id.* at 1070 ("[A] supplier may wrongly terminate a consumer agreement due to the consumer's race while mistakenly overlooking a statutory rule that prohibits discriminatory contract termination."); *see also*, El-Hallani v. Huntington Nat'l Bank, 623 F. App'x 730, 732–739 (6th Cir. 2015) (explaining that the plaintiffs alleged a bank closed their accounts because of their race, in violation of 42 U.S.C. §§ 1981 and 1982, and the Michigan Elliot–Larsen Civil Rights Act; the court ruled that the plaintiffs pleaded factual content that allowed the court to draw the reasonable inference that the defendant was liable for the misconduct alleged).

58. Benoliel & Becher, *supra* note 49, at 1064.

59. *See, e.g.*, Yonathan A. Arbel & Roy Shapira, *Consumer Activism: From the Informed Minority to the Crusading Minority*, 69 DEPAUL L. REV. 233, 259 (2020) (consumers' ability to file lawsuits against sellers "impose direct costs on firms, and the threat of these costs can deter seller misbehavior ex ante").

they can still influence firms' reputations and discipline them to avoid breaching a contract.[60]

To begin with, the mere *filing* of a breach of contract lawsuit may generate reputational costs to the breaching supplier. The willingness of an individual consumer to invest resources and time in filing a lawsuit may signal to other consumers that the probability that the supplier breached the contract is not trivial.[61] The complaint makes the risk of a breach more salient and vivid. As a result, potential consumers, who are informed about the lawsuit filing, may infer from the filing that the supplier might undervalue its clients' legal rights.[62] Consequently, potential consumers may avoid transacting with this supplier.[63] For similar reasons, existing consumers may also reduce their interactions with the supplier due to the lawsuit filing, thereby increasing the supplier's reputational costs.[64] Existing consumers may also closely examine their own interactions and contracts with the firm, and potentially file a suit of their own or initiate (or join) a class action. Here too, the same rationale pertains to other forms of consumer activism. Consumer complaints, negative online reviews, and

60. *See, e.g.*, Shmuel I. Becher & Tal Z. Zarsky, *Online Consumer Contracts: No One Reads, but Does Anyone Care?*, 12 JERUSALEM REV. LEGAL STUDIES 105, 109–110 (2015) (discussing the importance of press and media interest in consumer form contracts); Arbel & Shapira, *supra* note 35 (discussing the potential role of active and vocal consumers); Jeff Sovern, *Six Scandals: Why We Need Consumer Protection Laws Instead of Just Markets,* 11 MICH. BUS. & ENTREPRENEURIAL L. REV. 1 (2021) (examining famous incidents in which companies mistreated consumers and concluding that markets and reputation are not a consumer protection panacea).

61. *Cf.* David W. Prince & Paul H. Rubin, *The Effect of Product Liability Litigation on the Value of Firms*, 4 AM. L. & ECON. REV. 44, 57 (2002) ("[W]e can expect a negative impact on the value of a firm surrounding the filing of a lawsuit. This is in part because the filing of a lawsuit indicates that plaintiff attorneys believe that a favorable verdict is sufficiently probable to justify what is often a substantial investment.").

62. *Cf.* Assaf Jacob & Roy Shapira, *An Information-Production Theory of Liability Rules*, 89 U. CHI. L. REV., 1113, 1119 (2022) ("Upon hearing the bad news [about a company], stakeholders may infer that the company's 'type' is worse than they previously thought; for example, they may infer that the company does not invest enough in the quality or safety of its products.").

63. Prince & Rubin, *supra* note 61, at 51 ("[L]awsuits that are damaging to a firm's brand name are likely to be associated with reputation costs due to lower quasi-rents from future sales [to consumers]."); Jacob & Shapira, *supra* note 62, at 1119 ("[B]ad news about the company may lead to diminished future business opportunities.").

64. For a similar argument in a different context, *see* Uri Benoliel, *Reputation Life Cycle: The Case of Franchising*, 13 CHAP. L. REV. 1, 6 (2009) ("[A] franchisor who terminates the contract without good cause will encounter difficulties in retaining its other franchisees.").

critical mass media coverage create information flows that reach other prospective consumers.[65] These information flows make it more likely that other consumers will exercise vigilance and reduce their willingness to engage with the firm that breached its contract.

Information flows are not limited to consumers, of course. Investors, too, may be deterred by negative information flows and suits against a company.[66] Investors may be concerned about the potential negative economic implications of the breach of contract lawsuit (e.g., fewer future sales) or the negative publicity.[67] Investors may also infer from the lawsuit or the negative publicity that the supplier disrespects consumer rights.[68] Such disrespect may expose the supplier to additional breach of contract lawsuits by consumers and further consumer complaints, ultimately harming its investors.

Empirical evidence, albeit in a different context, indicates that the mere filing of a consumer lawsuit against a supplier can cause reputational harm to the supplier. For example, Professors David Prince and Paul Rubin examined whether firms in the automobile and pharmaceutical industries suffer reputational costs due to product liability lawsuits.[69] They found that firms facing lawsuits for their products suffer significant capital market losses.[70] The

65. *See generally* Shmuel I. Becher & Tal Z. Zarsky, *E-Contract Doctrine 2.0: Standard Form Contracting in the Age of Online User Participation*, 14 MICH. TELECOMM. & TECH. L. REV. 303 (2008) (explaining the potential power of online information flows to discipline firms). These information flows have their own limits, of course. *See, e.g.*, Yonathan A. Arbel, *Reputation Failure: The Limits of Market Discipline in Consumer Markets*, 54 WAKE FOREST L. REV. 1239, 1253 (2019) (explaining that reputational information can be costly to obtain, noisy, distorted, or ineffectual).

66. *Cf.* James M. Rice, *The Defensive Patent Playbook*, 30 BERKELEY TECH. L.J. 725, 750 (2015) ("A litigation risk . . . deters some investors . . . ."); Roy Shapira, *Mandatory Arbitration and the Market for Reputation*, 99 B.U. L. REV. 873, 885 (2019) ("Investors hearing about a corporate governance scandal will start demanding higher returns for their investment.").

67. Rice, *supra* note 66, at 750 ("A litigation risk . . . deters some investors who see the exposure as a limit to potential revenue.").

68. *Cf.* Jacob & Shapira, *supra* note 62, at 1119 ("Upon hearing the bad news [about a company], stakeholders may infer that the company's type is worse than they previously thought; for example, they may infer that the company does not invest enough in the quality or safety of its products.").

69. Prince & Rubin, *supra* note 61, at 45 (2002) (one of the study's goals was "to determine whether firms suffer reputation costs as a result of lawsuits.").

70. *Id.* at 71 (indicating that firms facing lawsuits for their products suffer capital market losses approximately equal to a worst-case scenario associated with the litigation).

study found that the values of automobile firms facing lawsuits fall anywhere from $276.45 million to $499.22 million.[71]

The reputational costs caused to the supplier by the mere filing of a breach of contract lawsuit may be increased by the legal process that follows such filing. The process may reveal negative information about the supplier's practices that the firm had previously concealed from the public.[72] Particularly, documents exposed by the supplier during the discovery stage, which can be made public or published, or behaviors and norms exposed during testimony in open court, may reveal illegal internal practices the supplier exercises.[73]

The reputational costs to the supplier, caused by information flows or a breach of contract lawsuit and the legal process that follows, may be intensified by the ultimate legal *outcome* of the lawsuit. The final decision by the court may highlight patterns of contractual misbehavior by breaching suppliers.[74] In addition, the public is likely to treat a court decision, provided by a neutral judge, as trustworthy.[75] The same is often true about media coverage. Accordingly, the firm risks further reputational costs once these negative information flows and judicial decisions reach the public. Thus, reputational costs can pressure firms into changing policies or avoiding breaching contracts with consumers.

2. Hidden Contracts Hinder Consumer Activism

We have seen that firms may be motivated to breach contracts inefficiently, harming consumers. We have also seen how

71. *Id.* at 61 ("[T]he firms facing the lawsuits fall in value anywhere from $276.45 million . . . to $499.22 million.").

72. Shapira, *supra* note 66, at 887 ("[L]itigation helps market players by uncovering new pieces of information on the corporate misconduct in question.").

73. *Id.* at 888 ("[L]egal documents that come out during pleading, discovery, or trial help not just by drawing outside observers' attention to a misbehavior they were not aware of, but also by adding detail and analysis on how things happened."); Roy Shapira, *Reputation through Litigation: How the Legal System Shapes Behavior by Producing Information*, 91 WASH. L. REV. 1193, 1214 (2016) ("[I]ntra-company emails being revealed only during discovery [may expose] exactly what top managers knew and when they knew it.").

74. Shapira, *supra* note 66, at 888 ("Judicial opinions are good at flashing out patterns of misbehavior . . . .").

75. *Id.* ("Judicial opinions are normally considered disinterested and fair . . . ."); Shapira, *Reputation through Litigation*, *supra* note 73, at 1232 ("When well-respected judges put their name on a certain version of the events, stakeholders are more likely to update their beliefs based on it.").

consumer activism, and especially the possibility of a lawsuit, mitigates firms' appetites to breach their contracts. However, while lawsuits may serve as a deterrence mechanism against breaches, hidden contracts erode the potential of this mechanism. Slightly restated, hidden contracts can discourage consumers from complaining and filing a breach of contract lawsuit against breaching firms.

Consumer contracts contain many provisions that interact with consumers' willingness to complain, share grievances, and litigate. To begin with, consumers are likely to feel more comfortable and motivated to confront the firm, complain against it, and file a breach of contract lawsuit where they are confident that the firm violated their rights. But it is hard for consumers to know whether firms breach their contractual obligations if they do not know what the relevant contracts that govern the parties' relationships says.

Moreover, to assess the expected benefits of potential litigation, consumers (or their lawyers) must evaluate, among other things, the probability of winning a lawsuit.[76] This evaluation requires reviewing the content of the contracts they entered.[77] For example, consumers may wish to know the states whose laws govern their cases. Some states have stronger consumer protection laws than others,[78] affecting consumers' probability of winning at trial.[79]

76. *Cf.* STEVEN SHAVELL, FOUNDATIONS OF ECONOMIC ANALYSIS OF LAW 390 (2004) ("[S]uit is more likely . . . the greater the likelihood of winning at trial.").

77. *Breach of Contract Lawyers in Washington D.C.*, HKM, https://hkm.com/washingtondc/breach-of-contract (last visited Sept. 2, 2023) ("The first thing your lawyer will do [before filing a breach of contract lawsuit] is review the contract so that they can understand each party's terms and obligations."); Jaclyn Wishnia, LEGALMATCH, *Breach of Contract Lawsuit: Suing for Breach of Contract*, https://www.legalmatch.com/law-library/article/breach-of-contract.html (last visited Oct. 3, 2023) ("Before filing a breach of contract claim, it is important to review the contract . . . .").

78. A prime example is California, which is known to have relatively strong consumer protection laws. *See, e.g.*, Wershaba v. Apple Computer, Inc., 110 Cal. Rptr. 2d 145, 160 (Cal. Ct. App. 2001) ("California's consumer protection laws are among the strongest in the country . . . ."). At the same time, Alabama has a history of relatively weak consumer protection regulation. *See, e.g.*, Melissa Briggs Hutchens, *At What Costs?: When Consumers Cannot Afford the Costs of Arbitration in Alabama*, 53 ALA. L. REV. 599, 601 (2002) (Alabama has a "history of weak consumer protection laws . . . .").

79. *Bach Talk: In Praise of Contingency Fees*, UNITED POLICYHOLDERS, https://uphelp.org/bach-talk-in-praise-of-contingency-fees (last visited Sept. 2, 2023) ("In [states with weak consumer protection laws], attorneys can't or won't work on a contingency fee because it will reduce the policy benefits the consumer has been deprived of to the point where it doesn't make economic sense to sue.").

Furthermore, consumers (or their advocates) may wish to access their contracts to check whether they include any time bars for filing a lawsuit, which may impact their chances of winning.[80]

The same analysis applies to consumers who wish to criticize a firm's behavior online, arbitrate cases, or complain to consumer organizations, enforcement agencies, or the media. People perceive legal rights as worth protecting; after all, this is why the law grants such rights. Thus, a consumer complaint based on a breach of contract is more likely to gain prominence, attract attention, prompt empathy, and stimulate action.[81]

On top of that, to assess the expected *benefits* of a lawsuit, consumers must also evaluate the magnitude of their gains, assuming they win a case.[82] And, once again, in order to estimate the expected gains, consumers must evaluate the contracts they signed. For example, consumers may wish to know whether their contracts include a limitation-of-damages clause, and if so, the scope of that clause (i.e., how much consumers could recover if they prevail).[83] Consumers may also want to examine whether the contract requires the supplier to reimburse them for their legal fees if the consumers win at trial.[84]

80. For instance, according to Tripadvisor's sign-up contract, a consumer is obliged to file a lawsuit against Tripadvisor within two years from the date on which the cause of action arose. *See Tripadvisor Terms, Conditions and Notices*, TRIPADVISOR, https://tripadvisor.mediaroom.com/us-terms-of-use#OLE_LINK23 (last visited Sept. 2, 2023) ("[Y]ou agree that you will bring any claim or cause of action arising from or relating to your access or use of the Services within two (2) years from the date on which such claim or action arose . . . .").

81. People generally overvalue formal aspects of contract law. *See* Tess Wilkinson-Ryan & David A. Hoffman, *The Common Sense of Contract Formation*, 67 STAN. L. REV. 1269 (2015) (elucidating how laypeople attribute excessive significance to formal facets of contract law).

82. STEVEN SHAVELL, FOUNDATIONS OF ECONOMIC ANALYSIS OF LAW 390 (2004) ("The plaintiff's expected benefits from suit involve possible . . . gains from trial.").

83. A limitation-of-damages clause may cap the scope of damages that a plaintiff is entitled to receive. *See, e.g.*, Gabrielle Nater-Bass & Stefanie Pfisterer, *Contractual Limitations on Damages*, GLOB. ARB. REV., (Dec. 19, 2022), https://globalarbitrationreview.com/guide/the-guide-damages-in-international-arbitration/5th-edition/article/contractual-limitations-damages ("Contractual limitations on damages are agreements whereby the parties limit or exclude the availability of damages that would otherwise be available under statutory law.").

84. For example, according to the sign-up contract of Alamy.com, "[t]he prevailing party will be entitled to recover its reasonable legal costs relating to that aspect of its claim or defense on which it prevails . . . ." *See Terms and Conditions, art. 18.10*, ALAMY, https://www.alamy.com/terms/us.aspx#Miscellaneous-terms (last visited Sept. 2, 2023).

Similarly, when it comes to initiating litigation, consumers may be interested in examining the potential *costs* of the lawsuit.[85] The lower the expected costs, the higher the chances that consumers will file suit.[86] To assess the expected costs, consumers must, once again, analyze the contract they entered into with the supplier. Here, consumers may examine whether the contract requires them to file a lawsuit in a specific jurisdiction, which may be a distant state or country, thereby exposing consumers to significant travelling and lodging costs during the lawsuit.[87] Consumers may also seek to check whether the contract requires them to file a lawsuit only with an arbitration organization, which can impact the costs of the legal procedure.[88] Moreover, consumers may examine whether the contract requires them to reimburse the supplier, if they lose at trial, for its attorney's fees, which are often significant.[89]

Hidden contracts hamper the ability of consumers to become familiar with the relevant contract provisions, which can be paramount to a decision about whether to complain or initiate litigation. Hidden contracts, by their nature, are not accessible to consumers after they are signed.[90] Hence, hidden contracts entail legal uncertainty. Under such contractual ambiguity, consumers may find it difficult, if not impossible, to make an informed decision on how to handle their grievances or disputes. Put simply, we hypothesize that hidden contracts make consumers less likely

85. SHAVELL, *supra* note 82.

86. SHAVELL, *supra* note 82 ("[S]uit is more likely the lower the cost of suit . . . .").

87. Tanya J. Monestier, *Forum Selection Clauses and Consumer Contracts in Canada*, 36 B.U. INT'L L.J. 177, 210 & n.190 (2018) (The consumer's costs of litigation outside her home jurisdiction may include "travel and lodging, and costs associated with being temporarily out of work.").

88. Some contracts provide that only the firm pays for arbitration processes, while others do not. At the same time, arbitration can also entail greater costs to consumers. *See, e.g.*, Joe Valenti, *The Case Against Mandatory Consumer Arbitration Clauses*, CAP, (Aug. 2, 2016), https://www.americanprogress.org/article/the-case-against-mandatory-consumer-arbitration-clauses ("[A]rbitration can be a more expensive and time-consuming resolution that makes it harder for victims to pursue their claims.").

89. Westlaw Classic, *General Contract Clauses: Litigation Costs and Expenses*, Practical Law Standard Clauses 3-540-2608 ("Attorneys' fees are typically the largest component of the cost of pursuing or defending litigation."); *see also* Edward L. Rubin, *Trial by Battle. Trial by Argument.*, 56 ARK. L. REV. 261, 288 (2003) (noting the exorbitant cost of litigation).

90. *See* discussion *supra* Section I.A.

to file a lawsuit or otherwise be vocal and assertive and enforce their rights.[91]

Empirical studies implicitly indicate that legal uncertainty about consumers' rights may reduce consumers' willingness to initiate litigation against suppliers. In this context, Professor Ronald Tipper analyzed interviews of a nationally representative sample of more than 950 consumers.[92] Among other things, the respondents were asked which, if any, third-party actions, including legal actions, they had taken against a supplier.[93] The results of the study show that consumers with less knowledge of their legal rights tended to seek third-party redress less frequently than consumers with more knowledge about these rights.[94]

Moreover, hidden contracts also present practical challenges to consumers who consider bringing a breach of contract lawsuit against suppliers. Under the law of several states, a plaintiff suing for violation of a written contract must attach that contract to the complaint[95] or state its substance.[96] Naturally, hidden contracts can prevent consumers from attaching a contract to a lawsuit or effectively stating the contract's substance, as required by law.

---

91. *Cf.* Michael L. Ursic, *A Model of the Consumer Decision to Seek Legal Redress*, 19 J. CONSUMER AFFS. 20, 26 (1985) ("If a person feels that success in court is probable, he or she is more likely to take action than a person who does not feel that winning in court is probable."); Roger Van den Bergh & Louis Visscher, *The Preventive Function of Collective Actions for Damages in Consumer Law*, 1 ERASMUS L. REV. 5, 14 (2008) ("It is also very difficult for consumers to assess whether manufacturers have obeyed safety regulations. Due to this information asymmetry, consumers may not start a lawsuit . . . .").

92. Ronald H. Tipper, *Characteristics of Consumers Who Seek Third Party Redress*, 43 CONSUMER INTS. ANN. 222, 223 (1997) ("The data set was based on telephone interviews among a nationally representative sample of 957 adults . . . .").

93. *Id*. ("The respondents were asked which (if any) third party action was taken.").

94. *Id*. at 225 ("Consumers with more knowledge of consumer rights . . . tended to seek some type of third-party redress more than their counterparts.").

95. *See, e.g.*, Ramirez v. Palisades Collection LLC, No. CIV.A.07-C-3840, 2008 WL 2512679, at *3 (N.D. Ill. June 23, 2008) ("Illinois law requires that a plaintiff suing for violation of a written contract must attach that contract to its complaint . . . ."); Strategic Mktg., Inc. v. Great Blue Heron Software, No. 15-CIV-80032, 2015 WL 11438209, at *6 (S.D. Fla. May 12, 2015) (Florida law "expressly require[s] the attachment of a contract to a pleading where the contract is material to the pleadings . . . ."); Harleysville Lakes States Ins. Co. v. Mason Ins. Agency, Inc., No. 255195, 2005 WL 2323814, at *2 (Mich. Ct. App. Sept. 22, 2005) ("With certain exceptions, a party suing on a contract must attach that contract to its complaint.").

96. *See, e.g.*, Target Nat'l Bank v. Kilbride, No. 2009-4291, 2010 WL 1435304 (Pa. C.P. Feb. 5, 2010) ("If the writing is not available to the pleader, the pleader may . . . state the substance of the writing in the pleading.").

In sum, firms that employ hidden contracts blur the tracks of the contracts after the consumers agree to them.[97] As a result, hidden contracts create legal uncertainty for consumers, precluding them from properly assessing their situation, their ability to complain, the legitimacy of the complaint, and the potential costs and benefits of suing the supplier.[98] Hidden contracts can prevent consumers from actively confronting the firm and informing the public about firms' misbehavior and inefficient breach of contracts.[99] Consequently, hidden contracts shield suppliers from the reputational and legal costs of their illegal behavior. Therefore, hidden contracts increase the risk that suppliers will breach their contracts inefficiently, harming consumers and society more generally.

II. THE EMPIRICAL TEST

Given the social costs of hidden contracts,[100] an important empirical question arises: Do mega websites, used by billions of consumers worldwide, utilize practices that facilitate the emergence of undesirable hidden contracts? This Part tackles that question. Section A describes our sample of 100 popular websites. Thereafter, section B explains our methodology in analyzing those websites and their contracts. Finally, section C details our findings: that hidden contracts are a prevalent problem.

A. Data

This Article's sample contains 100 of the most popularly used websites in the United States that meet three conditions: first, the website allows consumers to sign up to their services, normally via a button titled "sign up"; second, during the sign-up process, the website informs consumers that by signing up to the site, consumers agree to its standard form contract, normally titled "terms of use," "terms of service" or "terms and conditions;" third,

97. *Supra* Section I.A.

98. *See supra* text accompanying note 88.

99. *See supra* text accompanying notes 89–93. This is not to argue, of course, that hidden contracts negate *any* kind of consumer complaints. Even without access to their contracts, consumers can still criticize, for example, defective products or bad customer service.

100. *See* discussion *supra* Section I.B.

during the sign-up process, consumers are required to provide their contact information, such as their email address.[101]

To identify the most popular U.S. websites, we used DataForSEO's list of top 1,000 websites.[102] Since not all the popular websites met the study's three conditions, the 100th website in our sample is ranked 159 in popularity according to DataForSEO. Appendix A lists these 100 highly popular sites, which include, inter alia, Google, Spotify, YouTube, Etsy, Facebook, Instagram, eBay, LinkedIn and Amazon. Overall, the websites that constitute this Article's sample belong to highly heterogeneous categories. These include hotels and accommodations; computers, electronics, and technology; e-commerce and shopping; news and media; social media; search engines; streaming and online TV; dictionaries and encyclopedias; jobs and career; music; video games; science and education; visual arts and design; and real estate.[103]

B. Methodology

To test the frequency with which firms apply practices that facilitate the creation of hidden contracts, we took the following independent steps. First, for each sample website, we examined whether it provides consumers a copy of the contract they agreed to while signing up for the website. For that purpose, we provided an email address to the website during the sign-up process. Seven days later, we examined whether the email inbox or spam folders contained an email from the website with a copy of the sign-up agreement.

Second, for each sample website, we tested whether the site removed any contract terms it had amended from its public sphere. To that end, we initially had to identify those websites that amended their terms. We identified these websites by employing two indicators.

101. These conditions were set, since this empirical study aims to examine whether mega sites transparently send to consumers a copy of their standard form agreements, after consumers agree to said contracts.

102. *See Top 1000 Websites by Ranking Keywords*, DATAFORSEO (Aug. 20, 2022), https://dataforseo.com/top-1000-websites.

103. The website categories were identified using the SimilarWeb search engine. *See* https://www.similarweb.com (last visited Sept. 4, 2023).

1. The First Indicator of Amendments to the Contract

For each sample website, we first examined whether its terms of use webpage included a statement that implies that previous terms were amended. For example, Facebook's current terms of service page includes the following statement: "Date of Last Revision: Sept. 19, 2022."[104] Likewise, Amazon's current terms include the statement "Last updated: September 14, 2022."[105] These statements imply that Facebook and Amazon amended a previous version of their contract terms.

2. The Second Indicator of Amendments

For websites that did not include such a statement, we tested whether the website contained a statement about the effective date of the current terms. When we found such a statement in the terms, we checked—via the Wayback Machine online tool—whether the terms webpage existed before the effective date of the current terms.[106] If the result of this test was positive, we concluded that the current contract terms of the website modified a previous version. For instance, LinkedIn's current terms include a statement that their effective date is February 1, 2022.[107] However, according to Wayback Machine, the terms webpage already existed on April 29, 2013.[108] Therefore, in this and similar cases, we concluded that LinkedIn's current terms modify a previous version of the terms. By employing these two indicators, we established a list of websites that amended their terms. Diagram 1 below visually illustrates the major steps of the inquiry described above.

104. *Terms of Service*, FACEBOOK (Sept. 19, 2022), https://www.facebook.com/terms.php.

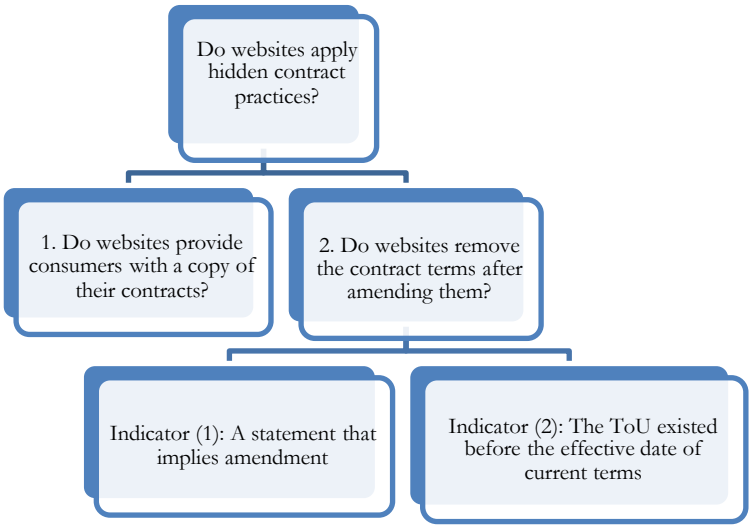
105. *Conditions of Use*, AMAZON (Sept. 14, 2022), https://www.amazon.com/gp/help/customer/display.html?nodeId=GLSBYFE9MGKKQXXM.

106. The Wayback Machine tool is available at https://archive.org/web (last visited Sept. 2, 2023).

107. *User Agreement*, LINKEDIN (Feb. 1, 2022), https://www.linkedin.com/legal/user-agreement.

108. *See* LinkedIn User Agreement, INTERNET ARCHIVE WAYBACK MACHINE, https://web.archive.org/web/20220000000000*/https://www.linkedin.com/legal/user-agreement (last visited Sept. 2, 2023).

Diagram 1. Analysis Inquiries



At this point, we moved to the next stage and examined whether the identified websites that amended their terms maintained a copy of pre-amendment versions of the terms on their website. For that purpose, we checked whether the terms webpage includes a link to either pre-amendment versions or an archive that includes these versions. For example, Yelp amended its contractual terms of use on December 13, 2019. Helpfully, Yelp's terms webpage includes a link to a digital archive of pre-amendment versions of these terms.[109] Conversely, Walmart.com amended its terms on June 19, 2023, but Walmart's terms webpage does not include a link to pre-amendment versions of the terms or an archive of previous versions.[110]

C. Results

The results of this study indicate that firms routinely apply practices that facilitate the emergence of hidden contracts. First, out

109. *See Terms of Service*, YELP, (DEC. 19, 2019), https://terms.yelp.com/tos/en_us/20200101_en_us (noting "to review the previous terms, please click").

110. *See Terms of Use*, WALMART, (June 19, 2023), https://www.walmart.com/help/article/walmart-com-terms-of-use/3b75080af40340d6bbd596f116fae5a0 ("The 'Last Updated' legend above indicates when these Terms of Use were last changed.").

of 100 sample websites, none sent consumers a copy of the standard form agreement to which the consumers agreed. Specifically, none of the websites email consumers either a PDF file with their sign-up agreement or the text of the contract.

Next, our findings indicate that most sample websites (82%; n=82) amended their original contract terms. As to the first indicator, the terms of 65 websites included a statement that implies that previous terms were amended (e.g., an announcement about the date of the last revision). Regarding the second indicator, among the remaining 35 websites, 20 included a statement about the effective date of terms. Of these 20 websites, the terms of use of 17 webpages existed earlier than the effective date presented in the current terms.[111] This implies that an older version of the terms was amended (by the current terms). Table 1 below summarizes these findings.

Table 1. Amendment indicators (out of 100 sample websites)

Indicator	Frequency
Implicit Statement	65
Effective Date	17
Total	82

Conspicuously, out of the 82 websites that amended their original contracts, the majority (80.5%; n=66) removed the pre-amended version of their contracts from their public spheres. Particularly, these websites failed to include on their terms webpages a link to the pre-amendment version of the terms or to an archive of previous versions.

Notably, our finding that 80.5% of the websites that amended their contracts removed the pre-amended versions from their site is conservative. To begin with, our two indicators of contract

111. The authors used the Wayback Machine online tool to examine whether the terms of use webpage existed earlier than the effective date of the terms. *See also supra* Section II.B.

amendment rely on what websites report and what we can find via the Wayback Machine. However, it is possible that firms amended their contracts non-transparently and did not indicate this fact on their websites. It is reasonable to suspect that some firms may not announce they modified the terms (indicator 1) or the terms' effective dates (indicator 2)—especially when such modifications harm consumers.[112] Secondly, and as we explain below, the Wayback Machine is an incomplete tool: it does not encompass all previous web pages, and firms can opt out of being included in it.[113]

Importantly, our empirical examination also revealed that—even among the minority of 16 websites that amended their contracts and published a purported archive with pre-amendment versions—the archive is typically incomplete. Particularly, 13 of these 16 websites (i.e., 81.25%) do not have full archives. In other words, only 3 of the 82 sampled websites that changed their contract terms included apparently adequate archives.[114]

The archives that do exist lack in various ways. In 5 websites, the oldest version of the contract in the archive does not seem to be the actual oldest version. To illustrate, according to the language of the oldest contract version published in Quora.com's archive, that version was last updated on "December 18, 2017."[115] However, Quora.com's archive does not include the version that was in effect before that update.[116]

Likewise, in 3 additional websites, the oldest version of the terms in the archive does not seem to be the actual oldest version of the terms, as reflected by the internet archive tool, the Wayback Machine. For instance, Houzz.com's terms of use webpage includes

112. *See* discussion *supra* Section II.B (explaining this study's methodology and elaborating on these two indicators). Theoretically, one may argue that perhaps the firms that changed their terms without using these two indicators did include an archive. To negate this possibility, we checked these websites and found that none of them included archives of previous contract versions.

113. *See* discussion *infra* Section III.C (explaining that not all previous versions of consumer form contracts are available in internet archives).

114. For example, Twitter's archive apparently includes all the versions of their terms (from version number one to seventeen). *See Previous Terms of Service,* TWITTER, https://twitter.com/en/tos/previous (last visited Sept. 2, 2023).

115. *Terms of Service – Prior Versions*, QUORA https://www.quora.com/about/tos_archive (last visited Sept. 2, 2023).

116. *Id*.

a link to a prior version of the terms,[117] which was effective on January 1, 2020.[118] However, Houzz.com's terms of use webpage does not include a link to any version of the terms that was effective before that date. This is so despite Houzz.com's terms of use webpage having existed much earlier than January 1, 2020, according to the Wayback Machine.[119] Similarly, another site, AliExpress, included the full text of only one older version of its terms, stating that it was "effective as of April 30, 2021."[120] Nevertheless, according to information provided by the site itself, it was founded much earlier, in 2010,[121] and ostensibly had terms of use soon thereafter.

Two other websites only mentioned historical changes made in their contract terms instead of a well-organized and systematic archive that includes the full text of all previous contracts. To illustrate, Tumblr.com includes a link to previous versions of its terms.[122] This link contains the following promising statement: "You will find prior versions of our Terms of Service on GitHub, which will allow you to compare historical versions and see which terms have been updated."[123] However, the link leads to a non-consumer-friendly list of sub-links, each connecting to a presentation of historical changes made in the terms.[124] The screenshots below illustrate.

117. *Houzz Terms of Use*, HOUZZ, https://www.houzz.com/termsOfUse (last visited Sept. 2, 2023) ("Click here for the prior revision").

118. *Id.* ("Effective January 1, 2020").

119. *Houzz's Terms of Use*, INTERNET ARCHIVE WAYBACK MACHINE, https://web.archive.org/web/20220000000000*/https://www.houzz.com/termsOfUse (last visited Sept. 2, 2023) (Houzz's terms of use webpage existed at least from July 10, 2009.).

120. *AliExpress.com Terms of Use*, ALIEXPRESS (June 13, 2022), https://terms.alicdn.com/legal-agreement/terms/suit_bu1_aliexpress/suit_bu1_aliexpress202204182115_66077.html?spm=a2g0o.home.0.0.650c2145ZCwgp.

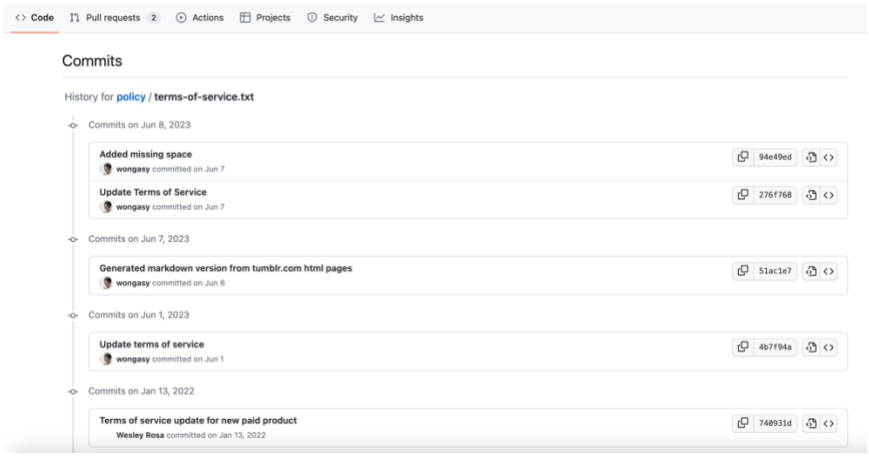
121. *Aliexpress.com, Website Biography, Domain Age, Reviews, Logo*, ONLINE BIOGRAPHY, https://onlinebiography.in/aliexpress (last visited Sept. 2, 2023) ("aliexpress.com was founded in the year 2010.").

122. *Terms of Service*, TUMBLR (June 7, 2023), https://www.tumblr.com/policy/en/terms-of-service.

123. *Id*.

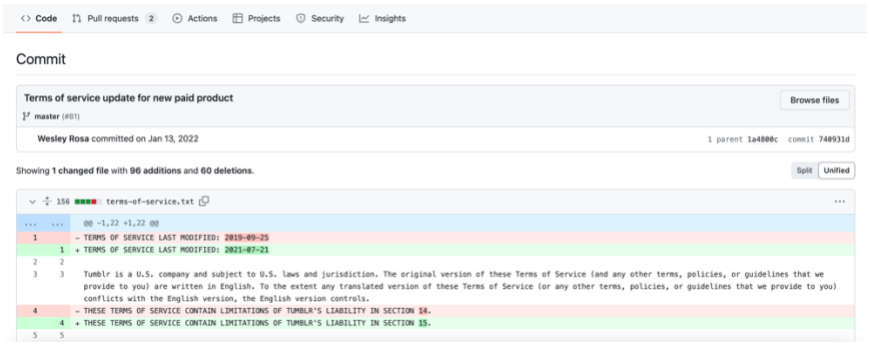
124. *Tumblr/Policy*, GITHUB, https://github.com/tumblr/policy/commits/master/terms-of-service.txt (last visited Sept. 2, 2023).

Image 1. Tumblr.com link to previous versions of its terms



Clicking on the last link above (titled "Terms of service update for new paid product") brings the user to the following awkward webpage.

Image 2. Tumblr.com's presentation of historical changes

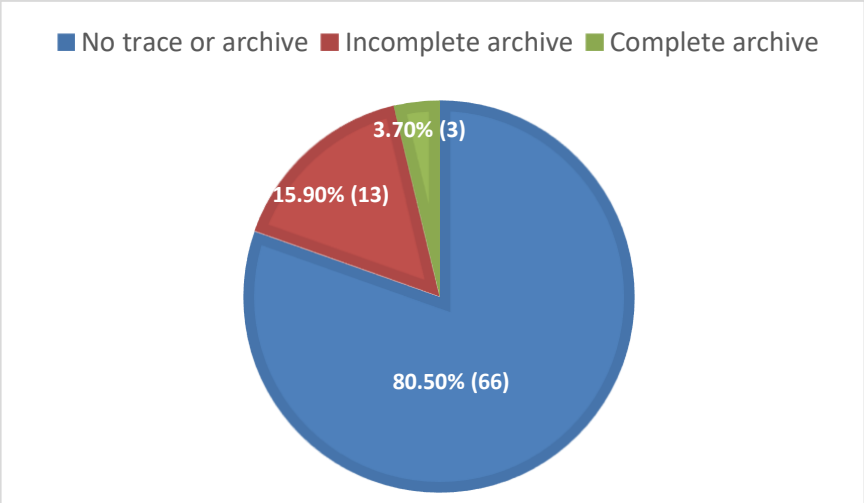


Another website, Github.com, did not present any older versions of the site's terms, although the site includes a link stating: "You can view all changes to these Terms in our Site Policy repository."[125] Lastly, on one website, Alibaba.com, the terms'

125. *GitHub Terms of Service*, GITHUB, https://docs.github.com/en/site-policy/github-terms/github-terms-of-service (last visited Sept. 2, 2023).

historical archive is only available in Chinese.[126] The graph below summarizes our findings regarding the 82 websites that amended their contractual terms.

Graph 1. Traces of previous versions (N=82)



III. POLICY RECOMMENDATIONS

Firms often apply practices that underlie hidden contracts.[127] Online contracting realities change the contracting ecosystem, allowing firms to constantly write over their terms and modify their contracts at virtually no cost. However, businesses consistently fail to send a copy of their standard form contracts to consumers,[128] and

126. The link "history rules" in Alibaba's terms of use webpage leads to an archive in Chinese. *See* ALIBABA.COM, https://rulechannel.alibaba.com/icbu#/rules?cId=-1 (last visited Sept. 4, 2023). In addition, Alibaba's terms of use webpage includes an English version of the full text of only one older version of the terms, "effective as of April 30, 2021," although the site was founded in 1999. *See Terms of Use*, ALIBABA.COM (July 6, 2021), https://rulechannel.alibaba.com/icbu?type=detail&ruleId=2041&cId=1307#/rule/detail?cId=1307&ruleId=2041. *See also Alibaba.com vs AliExpress: What are the Differences*, ALIBABA (Dec. 10, 2020), https://seller.alibaba.com/businessblogs/px53308i-alibabacom-vs-aliexpress-what-are-the-differences ("Alibaba.com was founded in 1999 . . . .").

127. *See* discussion *supra* Section II.C.

128. *See* discussion *supra* Section II.C.

when they amend their contracts, they frequently do not maintain the original contract versions on their websites.[129]

The implications of this practice are alarming. Hidden agreements hinder the ability of consumers (or their lawyers and consumer organizations) to analyze the contracts they agreed to.[130] As a result, consumers may be unable to accurately assess the potential costs and benefits of a breach of contract lawsuit against a breaching supplier.[131] Consequently, consumers may be deterred from complaining against the supplier or filing a breach of contract lawsuit against inefficient breaches of their standard form contracts.[132] This reality is undesirable because consumers' ability to air their complaints and initiate litigation is an important mechanism that empowers consumers and disciplines sellers.[133]

This Part concisely points to a few possible law and policy responses. Section A proposes introducing a contract transparency duty that will oblige firms to provide consumers with a copy of their contracts and maintain the versions of their contracts on their websites. Next, section B details ways to enforce this duty and enhance its effectiveness. It mainly suggests administrative enforcement and using injunctions, penalties, and fines. Thereafter, section C examines whether employing hidden contracts amounts to an unfair and deceptive practice under existing consumer law, highlighting the potential and obstacles of such a claim.

A. Transparency Duty

In many domains, the best way to solve a problem is to prevent it. The case of hidden contracts is no different. Given the social costs of hidden contracts, policymakers should consider imposing a *contract transparency duty* on firms. Such a duty, which should encompass a few operational aspects, will operate ex ante to discipline sellers.

First, under this duty, firms would be primarily required to provide consumers with their contracts following the formation of the agreement. However, we acknowledge that this is far from a

129. *See* discussion *supra* Section II.C.
130. *See* discussion *supra* Section I.B.
131. *See* discussion *supra* Section I.B.
132. *See* discussion *supra* Section I.B.
133. *See* discussion *supra* Section I.B.

magic bullet. Consumers accept a plethora of contracts during their lives, and one cannot reasonably expect all consumers to keep all of their contracts accessible all of the time.[134] For various legitimate reasons, consumers may lose the original copies of contracts provided by suppliers.

Accordingly, the second component of the proposed transparency duty would oblige suppliers to publish in their public spheres (e.g., their websites) all the versions of their standard form contracts, including all original and amended versions and the dates they were in force. This duty would allow consumers to easily review, at any stage, any previous versions of suppliers' contracts. Importantly, requiring firms to publish earlier versions of their agreements is not too burdensome. Businesses are likely to maintain previous versions of their contracts for internal uses, and if they do not currently retain such versions, they can easily adopt the practice.

One may argue that these two mechanisms may not suffice, since consumers may not know which of the previous versions applies in their individual cases. Moreover, some consumers—especially vulnerable consumers, such as the elderly, non-native English speakers, those who experience learning difficulties, and less-educated populations—may find it difficult to access and navigate online links containing legal documents. Accordingly, policymakers could add a third component to the proposed duty of transparency. Under this third component, firms would have to reproduce the original contract at any stage of the contractual relationship upon a consumer's request.

To be sure, a duty to reproduce a copy of the agreement upon the consumer's request may impose administrative costs on the supplier (e.g., time spent by employees). However, these costs are likely to be relatively low. Technological tools can make it quite easy for businesses to locate and provide the relevant contracts to those consumers who request them. Presumably, firms can easily track when a consumer joined their services and effortlessly link these points in time to the applicable contracts. That said, if policymakers conclude that the costs involved are considerable,

134. Paraphrasing Bob Marley, we opine that you might expect consumers to keep some contracts sometimes, but you can't assume they will keep all the contracts all the time. We intuit that many consumers would (often mindlessly) delete emails that contain their form contracts.

they can determine that the requesting consumer shall reimburse the supplier for reasonable administrative expenses. Charging an administrative/copy fee exists in other domains, such as requesting copies of medical records and account statements from a lender or creditor.

The contract transparency duty proposed in this Article is socially desirable. Importantly, this duty would help tackle the long-lasting challenge of consumer access to justice.[135] A transparency duty would assist consumers in better understanding their contractual legal rights once a supplier harms them. In doing so, the contract transparency duty may assist consumers in making informed decisions about complaining to third parties, sharing their experiences online, or filing breach of contract lawsuits against suppliers. Such consumer activism can expose suppliers to considerable reputational costs.[136] This exposure, in turn, may desirably deter suppliers, ex ante, from inefficiently breaching their contracts.

Furthermore, a duty of transparency would facilitate more accountability and, ideally, encourage a fairer and more balanced B2C environment. Online realities enhance firms' capability to manipulate consumers in many subtle ways, including utilizing nuanced design features (also known as. dark patterns).[137] Technological developments produce winners and losers. Transparency laws have the untapped potential to compel firms to use technology (cheap means of communication and webpages) and data (regarding dates of consumers' sign-ups and the contracts they accepted) to benefit consumers at a very low cost.

Remarkably, the Consumer Financial Protection Bureau (CFPB) requires credit card companies to make their agreements available

135. *See, e.g.*, Yehuda Adar & Shmuel I. Becher, *Ending the License to Exploit: Administrative Oversight of Consumer Contracts*, 62 B.C. L. REV. 2405, 2439–41 (2021) (reviewing the typical hurdles consumers face in enforcing their rights and litigating their cases).

136. *See* discussion *supra* Section I.B.2.

137. *See, e.g.*, Kerstin Bongard-Blanchy, Arianna Rossi, Salvador Rivas, Sophie Doublet, Vincent Koenig & Gabriele Lenzini, *"I am Definitely Manipulated, Even When I am Aware of It. It's Ridiculous!" – Dark Patterns from the End-User Perspective*, DESIGNING INTERACTIVE SYS. CONF. 2021 (June 28–July 2) (finding that dark patterns can manipulate and influence consumers even if consumers are aware of them).

on the CFPB's website, which contains prior agreement versions.[138] Recently, the CFPB also proposed creating a public registry of problematic standardized terms in nonbank customer agreements, citing concerns about lengthy one-sided contracts and the need to better monitor contracting practices.[139] According to this proposal, the CFPB would publish the information it collects in an accessible and centralized database online.[140] Notably, the CFPB's director highlighted the regulatory need to easily spot the use of harmful boilerplate, alongside the benefits of providing consumers with accessible information regarding form contracts.[141]

Moreover, a few states have already adopted a contract transparency duty similar, to some extent, to the duty proposed in this Article. To illustrate, some states require suppliers to provide consumers with a copy of the contract following its formation.[142] However, this duty often focuses only on a few limited types of transactions, including the sale of vacation time-sharing plans,[143] motor vehicles,[144] home improvement services,[145] insurance,[146]

138. *See Credit Card Agreement Database,* CFPB, https://www.consumerfinance.gov/credit-cards/agreements/?_gl=1*1y96w9v*_ga*NTI0Mzc1OTAxLjE2NjExODMyMTg.*_ga_DBYJL30CHS*MTY3MzMwOTg4Mi4xMS4xLjE2NzMzMDk5MDguMC4wLjA (last visited Sept. 2, 2023) ("The CFPB maintains a database of credit card agreements from hundreds of card issuers.").

139. *See CFPB Proposes Rule to Establish Public Registry of Terms and Conditions in Form Contracts that Claim to Waive or Limit Consumer Rights and Protections,* CFPB (Jan. 11, 2023), https://www.consumerfinance.gov/about-us/newsroom/cfpb-proposes-rule-to-establish-public-registry-of-terms-and-conditions-in-form-contracts-that-claim-to-waive-or-limit-consumer-rights-and-protections.

140. *See id.*

141. *See id.*

142. *See, e.g.*, DEL. CODE ANN. tit. 6, § 2825 (West) ("The seller shall deliver to the purchaser a fully executed copy of the contract . . . ."); OHIO REV. CODE ANN. § 4517.26 (West 2002) ("The seller, upon execution of the agreement or contract and before the delivery of the motor vehicle, shall deliver to the buyer a copy of the agreement or contract . . . .").

143. *See, e.g.*, tit. 6, § 2825 ("The seller shall deliver to the purchaser a fully executed copy of the contract . . . .").

144. *See, e.g.*, § 4517.26 ("The seller, upon execution of the agreement or contract and before the delivery of the motor vehicle, shall deliver to the buyer a copy of the agreement or contract . . . .").

145. *See, e.g.*, 73 PA. STAT. & CONS. ANN. § 517.7(c) (West 2014) ("A contractor or salesperson shall provide and deliver to the owner, without charge, a completed copy of the home improvement contract at the time the contract is executed . . . .").

146. *See, e.g.*, S.C. CODE ANN. § 38-63-210 (1976) ("Every insurer doing a life insurance business in the State shall deliver with each policy of insurance issued by it a copy of the application made by the insured so that the whole contract appears in the application and policy of insurance.").

health care plans,[147] manufactured or mobile homes,[148] and goods or services paid for in installments.[149] Similarly, in the banking industry, suppliers are required in some types of transactions to "(i) [p]ost and maintain the consumer's agreement on [their] Web site[s]; or (ii) Promptly provide a copy of the consumer's agreement to the consumer upon the consumer's request . . . ."[150] However, since the social costs of hidden contracts are not confined to a limited set of relatively high-value or high stakes transactions, policymakers should consider expanding this obligation and imposing a *general* contract transparency duty that governs all mass consumer contracts.

Interestingly, a few jurisdictions have adopted a general consumer contract transparency duty, not limited to a particular category of transactions. For example, under the California Consumer Contract Awareness Act of 1990, a supplier is generally required to deliver a copy of its contract to the consumer.[151] Although we welcome this approach, the Californian model can be improved. First, providing the contract to the consumer after entering it is merely one (admittedly, the first) component of transparency. As explained above, suppliers should be obliged to maintain all the versions of their consumer contracts in their public spheres (for instance, their websites).[152] In addition, policymakers

147. *See, e.g.*, W. VA. CODE ANN. § 33-25-13(a) (West 1971) ("Every such corporation shall deliver to each subscriber to its health care plan a copy of the contract.").

148. *See, e.g.*, OHIO REV. CODE ANN. § 4781.24(a) (West 2010) ("The seller, upon execution of the contract and before the delivery of the manufactured or mobile home, shall deliver to the buyer a copy of the contract . . . .").

149. *See, e.g.*, FLA. STAT. ANN. § 520.34(c) (West 2003) ("The seller shall deliver to the buyer, or mail to the buyer at his or her address shown on the contract, a copy of the contract signed by the seller.").

150. 12 C.F.R. § 1005.19 (2019)

151. CAL. CIV. CODE § 1799.202(a) (West 1991) ("[A] seller shall deliver a copy of a consumer contract to the consumer at the time it is signed by the consumer . . . ."). *See also* TEX. BUS. & COM. CODE ANN. § 601.052(a) (West 2009) ("A merchant must provide a consumer with a complete receipt or copy of a contract pertaining to the consumer transaction at the time of its execution."); ME. REV. STAT. ANN. tit. 32, § 4662(1) (2004) (in consumer solicitation sales, "[t]he seller shall furnish a completely executed copy of the contract or agreement to the consumer immediately after the consumer signs the agreement or contract.").

152. *See* discussion *supra* Section III.A (proposing a contract transparency duty that obliges online firms to have previous versions of the consumer contracts available and accessible online).

could direct firms to send to the consumer, upon the latter's request, a copy of the contract.

B. A Duty that Bites: Adjudication and Administrative Enforcement

Introducing a transparency duty is the first positive step in tackling hidden contracts. Alongside this step, it is crucial to implement enforcement mechanisms and principles that incentivize firms to comply. In other words, policymakers should ensure that the conceptual idea of a transparency duty enjoys an operational and effective framework that deters firms from employing hidden contracts.

Indeed, an explicit and concrete remedy is another aspect lacking in the California Consumer Contract Awareness Act. According to the existing model's remedial scheme, a supplier that fails to comply with the duty to provide a copy of the agreement is liable to the consumer for any resultant actual damages suffered by the consumer.[153] However, as has been argued in this Article, a major problem created by hidden contracts is that they keep consumers in the dark about their contractual rights.[154] This legal obscurity may hinder consumers from being able to sue suppliers for the damages consumers suffer due to suppliers' failures to provide agreement copies. Thus, the legal process that would determine such damages may never begin.

Hence, regulators should back the proposed transparency duty with an explicit remedy that directly tackles the non-transparency that hidden contracts create. In this context, there are a few options to consider. First is the possibility of private enforcement. If a consumer successfully sues a seller who fails to comply with the contract transparency duty, the consumer should be entitled to an injunction that compels the seller to comply. Additionally, if the request for an injunction is deemed justified, the supplier should reimburse the consumer for all the reasonable costs involved (including, for example, loss of time and legal fees). To further incentivize consumers, lawmakers could also allow statutory

153. CAL. CIV. CODE § 1799.205 (West 1991) ("A seller who fails to comply with section 1799.202 is liable to the consumer for any actual damages suffered by the consumer as the result of that failure.").

154. *See* discussion *supra* Section I.B.

multiple and minimum damage awards that courts can grant successful plaintiffs.[155]

Such a scheme notwithstanding, relying only on consumers to enforce the law and discipline firms may be unwise, if not unfair. Even under a transparency duty, some consumers may not be aware of their rights, prefer not to spend their limited time reading consumer contracts, fear confronting firms, or lack the motivation or necessary resources to litigate their cases.[156] Furthermore, mandated arbitration clauses, which are part of many standard form contracts, considerably limit consumers' ability to bring their claims before the courts.[157] Therefore, another imperative path to consider is administrative enforcement.

In view of that, we suggest that administrative agencies be vested with the authority to impose a fine for violating the transparency duty. Certainly, there is a growing recognition of the potential of administrative enforcement tools to contribute to a more balanced B2C environment and increase firms' compliance. For example, the recent Directive (EU) 2019/2161 empowers European Union Member States "to decide on the administrative or judicial procedure for the application of penalties for infringements" of Council Directive 93/13/EEC on unfair terms in consumer contracts.[158] Directive 2019/2161 further acknowledges that

155. *Cf.* N.Y. GEN. OBLIG. LAW § 5-702 (2020) (requiring that consumer contracts employ plain language and providing for a penalty of actual damages plus $50).

156. *See, e.g.*, U.S. FED. TRADE COMM'N, CONSUMER FRAUD IN THE UNITED STATES: AN FTC SURVEY 80–81, 80 tbl.5-1 (2004) (finding that only less than one tenth of defrauded consumers complained to official sources).

157. *See, e.g.*, Imre Stephen Szalai, *The Prevalence of Consumer Arbitration Agreements by America's Top Companies*, 52 U.C. DAVIS L. REV. ONLINE 233, 234 (2019), https://lawreview.law.ucdavis.edu/online/vol52/52-online-Szalai.pdf ("[A]t least 826,537,000 consumer arbitration agreements were in force" in 2018.); Thomas H. Koenig & Michael L. Rustad, *Fundamentally Unfair: An Empirical Analysis of Social Media Arbitration Clauses*, 65 CASE W. RSRV. L. REV. 341, 341, 351 (2014) (defining forced arbitration clauses in consumer contracts as a U.S. phenomenon and finding that U.S.-based social media providers were about three times more likely to incorporate a mandatory arbitration clause in their terms of use than non-U.S.-based providers); U.S. CONSUMER FIN. PROT. BUREAU, *Section 2: How Prevalent Are Pre-dispute Arbitration Clauses and What Are Their Main Features?*, *in* ARBITRATION STUDY: REPORT TO CONGRESS, PURSUANT TO DODD-FRANK WALL STREET AND CONSUMER PROTECTION ACT § 1028(A), 25–27, § 2.3.6 & fig.6 (2015) (finding that 87.5% of the major wireless providers have arbitration obligations in their contracts).

158. *See* Directive (EU) 2019/2161 of the European Parliament and of the Council of 27 November 2019 amending Council Directive 93/13/EEC and Directives 98/6/EC, 2005/29/EC and 2011/83/EU of the European Parliament and of the Council as regards the better enforcement and modernization of Union consumer protection rules, art. 14.

administrative authorities could impose penalties against firms (in this case, firms that employ unfair contract terms).[159] A similar logic should apply to hidden contracts.

Administrative enforcement has multiple advantages, and administrative agencies can be more effective than courts at protecting consumers from harmful practices and exploitation.[160] First, administrative agencies can be proactive rather than passive and reactive. Unlike courts, administrative bodies do not depend on the initiative of private consumers. Second, whereas courts respond to individual cases with no ability to prioritize them, administrative enforcement authorities can channel public resources more efficiently. That is, they can adopt a macro perspective and focus on the most pressing, prescient, or harmful practices and firms.

Furthermore, unlike the judiciary, administrative bodies can engage in a dialogue with various stakeholders, such as consumers, consumer organizations, firms, and other agencies. Agencies are also less bounded in the remedies and solutions they may seek. For example, administrative authorities can reach a consensual agreement with the firm and accompany and monitor its implementation.

Furthermore, agencies operating at the federal level, such as the Federal Trade Commission (FTC) and the CFPB, have access to complex information and can consult with academics and experts. These agencies, and their counterparts on the state level, can develop industry-specific expertise, which elected judges may not possess. Additionally, administrative bodies address market-wide issues rather than responding on a case-by-case basis. This entails that administrative enforcement can be more coherent and much faster than judicial review. Overall, administrative bodies have greater institutional competence, better capability in responding to market dynamics, and higher capability to avoid political pressure. They can therefore supplement important private enforcement initiatives in valuable ways.

Particularly, administrative agencies should be able to make two key remedial responses when tackling hidden contracts. First,

159. *Id.*

160. The discussion regarding the advantages of administrative enforcement relies on Adar & Becher, *supra* note 135, at 2443–46.

they should be empowered to issue orders—such as mandatory injunctions and cease and desist notices—to ensure that firms comply with the transparency duty. Specifically, this authorization warrants that firms (1) provide consumers with their original contracts and (2) maintain an adequate online archive. Second, administrative agencies should be authorized to issue penalties, sanctioning noncompliance. These penalties should be sufficiently steep to deter firms from utilizing hidden contracts, and they should not be arbitrarily capped.[161] Of course, any administrative decision should be subject to judicial review or an appeal before an administrative tribunal.

Finally, the content and imposition of any such remedies should be publicly transparent and visible. For example, they could be posted on the agency's website. Such publicity would maximize the impact of these orders and sanctions. At the same time, enforcement agencies could also include examples of "best practice" archives on their websites. These positive examples will serve as role models and enhance the reputation that compliance entails.

A. Hidden Contracts as an Unfair Practice

Whereas a contract transparency duty does not currently exist, all states have enacted a consumer protection statute (also known as a UDAP law).[162] Though these statutes vary in scope and effect, they generally prohibit deceptive, unfair, unconscionable, or abusive practices.[163] Notably, State Attorneys General can bring enforcement actions under their states' UDAP laws.[164] Furthermore, some of these laws empower consumers to pursue

161. *See generally* CAROLYN CARTER, NAT'L CONSUMER L. CTR., CONSUMER PROTECTION IN THE STATES: A 50-STATE EVALUATION OF UNFAIR AND DECEPTIVE PRACTICES LAW 1, 44, 55–56, 58–60, 62–63 (2018) (referring to ceilings such as $1,000 on civil penalties as obstacles in the path of consumer protection enforcement).

162. *See* NAT'L CONSUMER L. CTR., *Unfair, Deceptive and Abusive Practices (UDAP)*, https://www.nclc.org/topic/unfair-deceptive-and-abusive-practices-udap (last visited Sept. 29, 2023) ("Every state has a consumer protection statute . . . .").

163. *See id.*

164. *See, e.g.*, CROWELL, *State UDAP Laws are a Treasure Trove of Civil Penalties for Price Gouging*, (Apr. 15, 2020), https://www.crowell.com/NewsEvents/AlertsNewsletters/all/State-UDAP-Laws-are-a-Treasure-Trove-of-Civil-Penalties-for-Price-Gouging ("State Attorneys General have statutory authority and power to enforce consumer protection laws, and many have warned that they will do so.").

private action against non-compliant businesses,[165] and many delegate rulemaking authority to the state consumer protection agencies.[166] The FTC and CFPB also have a broad statutory authority to prohibit unfair and deceptive practices.[167]

Overall, these statutes should play a central role in protecting consumers against dishonest businesses.[168] Thus, it is important to consider whether employing hidden contracts amounts to an unfair practice. Below we focus on what is an unfair practice under the FTC Act, which inspired much of state UDAP laws.[169] In fact, UDAP laws are frequently regarded as "little FTC Acts," despite occasionally deviating from that Act.[170] We narrow our attention to unfair practices because hidden contracts are less likely to satisfy the criteria for a deceptive practice according to the FTC Act.[171]

165. This is the case, for example, in Nebraska, Alaska, Maryland and New Mexico. *See id.* For analyzing the interplay between federal, state, and private enforcement of consumer protection laws, see Dee Pridgen, *The Dynamic Duo of Consumer Protection: State and Private Enforcement of Unfair and Deceptive Trade Practices Laws*, 81 ANTITRUST L.J. 911 (2017).

166. *See* David Berman, Note, *A Critique of Consumer Advocacy Against the Restatement of the Law of Consumer Contracts,* 54 COLUM. J.L. & SOC. PROBS. 49, 52, 85–89 (2020) (suggesting using UDAP laws to tackle unfair contract terms).

167. *See* Federal Trade Commission Act, 15 U.S.C. § 45(a)(1)–(2) (FTC's mandate to prevent "deceptive acts or practices in or affecting commerce"); Dodd-Frank Act, § 1031, 12 U.S.C. § 5531(a)–(b); Dodd-Frank Act, Pub. L. No. 111-203, 124 Stat. 1376, § 1012(a)(10) (July 21, 2010) (CFPB's mandate).

168. *See* CARTER, *supra* note 161, at 1 ("Unfair and Deceptive Acts and Practices (UDAP) laws should be the backbone of consumer protection in every state."). *See also* Pridgen, *supra* note 165, at 912 ("State consumer protection statutes, known as state UDAP laws or state 'little FTC acts,' provide a stronghold of effective consumer protection in the United States.").

169. *See, e.g.*, Pridgen, *supra* note 165, at 917 ("Thus, the state UDAP laws, by providing for private enforcement of state laws that mimic the language of the federal law, have the effect of enlisting private consumer plaintiffs in the FTC's efforts to stem unfair and deceptive trade practices. Indeed, further solidifying the tie between the FTC Act and the state consumer protection laws, most of the state UDAP laws contain a provision declaring that the state legislature intended that the state courts and government enforcers be guided by relevant interpretations of the FTC Act in applying their own state law." (footnotes omitted)).

170. *See, e.g.*, *id.* at 912 ("State consumer protection statutes [are] known as state UDAP laws or state 'little FTC acts' . . . ."); Henry N. Butler & Joshua D. Wright, *Are State Consumer Protection Acts Really Little-FTC Acts*?, 63 FLA . L. REV. 163, 163 (2011).

171. For a practice to be deceptive, it should be likely to mislead a consumer acting reasonably under the circumstances and be material. *See* FED. TRADE COMM'N, POLICY STATEMENT ON DECEPTION 2, 5 (Oct. 14, 1983). Examples of deceptive acts or practices include misleading price claims, bait-and-switch techniques—where sellers make an alluring yet insincere offer to sell a product or service while intending to switch the consumer from the advertised merchandise to a different product or service that better benefits the sellers (see

Fairness is a vague legal norm, and the law does not precisely define what constitutes an unfair trade practice.[172] Rather, case law gradually defines the boundaries of unfair (and deceptive) practices and illustrates their scope[173]—at times, pushing their limits.[174] Though a trade practice can be both deceptive and unfair, it need not be deceptive to be unfair. Under current law, an unfair practice should (a) "cause[] or [be] likely to cause substantial injury which is [(b)] not reasonably avoidable by consumers . . ., and [(c)] not outweighed by countervailing benefits to consumers or to competition."[175]

First, the injury should be substantial and not trivial or speculative.[176] Typically, substantial injury comes in the form of monetary harm.[177] Importantly, a practice that causes a small amount of harm to many consumers may meet the substantial injury threshold.[178] Notably, actual injury is not always required, and a considerable risk of concrete harm may suffice.[179] In our context, the financial injury to consumers stems from the fact that firms blur the traces of their previous contracts. Slightly restated, hidden contracts prevent consumers from knowing their rights and obligations. This, in turn, undermines consumers' ability to complain about the firm's behavior, air their dissatisfaction, and bring their cases before the courts.[180] Additionally, determining the presence of substantial injury may

FTC, Guides Against Bait Advertising, 16 C.F.R. § 238.0), or omitting material limitations or conditions. *See* FTC Guides for the Advertising of Warranties and Guarantees, 16 C.F.R. §§ 239.1, 239.3.

172. *See* FTC v. R.F. Keppel & Bro., 291 U.S. 304, 310–12 (1934) (The Federal Trade Commission Act's ban on unfair competition tactics "does not 'admit of precise definition . . . .'" (quoting FTC v. Raladam Co., 283 U.S. 643, 648 (1931))).

173. *See id.* at 312 ("the meaning and application of [the Federal Trade Commission Act's unfairness ban] must be arrived at by what this court elsewhere has called 'the gradual process of judicial inclusion and exclusion.'") (quoting *Raladam, supra* note 172, at 648).

174. *Cf.* Pridgen, *supra* note 165, at 921–22 ("Some states . . . were leaders in aggressively enforcing their state consumer protection laws and also in pushing the boundaries of the meaning of 'unfair' or 'deceptive' trade practices.").

175. 15 U.S.C. § 45(n).

176. FED. TRADE COMM'N, POLICY STATEMENT ON UNFAIRNESS (Dec. 7, 1980).

177. *Id.*

178. *Id.* n.12.

179. *Id.*

180. *See* discussion *supra* Section I.B.

include public policy considerations, such as greater transparency and a fairer market environment.[181]

Assuming the "substantial injury" requirement is satisfied, one should examine whether consumers can reasonably avoid the injury. We opine in section IV.A below that it is unrealistic to expect consumers to always keep track of their consumer form contracts.[182] Firms are the least-cost avoiders, as they opt to employ hidden contracts, and they can easily solve the problems they create by posting previous versions of their contracts online (at a negligible cost). Conversely, and as we will explain in more detail in section IV.A, consumers, who are often time-constrained and boundedly rational, cannot be realistically expected to combat hidden contract practices by diligently saving all the consumer contracts they ever agree to.[183]

We further note, in this context, that the unfairness analysis adopts a "reasonable consumer" standard, not the "perfect consumer" who exercises "perfectly rational behavior."[184] Essentially, "the question 'is not whether a consumer could have made a better choice'"[185] (theoretically, build a personal archive of form contracts). In short, the law merely expects consumers to take reasonable actions to avoid injury.[186] Whereas a cost-benefit analysis may lead even a perfectly rational consumer to avoid maintaining a personal archive of consumer contracts, the reasonable consumer should definitely not be expected to do so.

Interestingly, one path the FTC uses to assess whether consumers could have reasonably avoided the injury is to examine whether the practice interferes with consumer decision-making.[187]

181. The FTC policy narrowed the scope for making public policy arguments to substantiate unfairness, but it did not entirely close it. Hence, public policy interest can be used to substantiate the injury and meet the substantial injury test.

182. *See* discussion *infra* Section IV.A. (refuting the argument that consumers should have the responsibility to maintain a personal archive with all their form contracts).

183. *See infra* Section IV.A.

184. CAROLYN CARTER, NAT'L CONSUMER L. CTR., UNFAIR AND DECEPTIVE ACTS AND PRACTICES § 4.3.2.3.2 (2021) (citing CONSUMER FIN. PROT. BUREAU [CFPB], CFPB SUPERVISION AND EXAMINATION MANUAL, UDAAP, UDAAP 2, 6 (2020)).

185. *Id*. (quoting CFPB, *supra* note 184, at 2).

186. *See id.*

187. *See* Rory Van Loo, *Helping Buyers Beware: The Need for Supervision of Big Retail*, 163 U. PA. L. REV. 1311, 1374–75 (2015) ("'Most of the Commission's unfairness matters are brought . . . to halt some form of seller behavior that unreasonably creates or takes advantage

Consumers cannot reasonably avoid injury if the practice at stake inhibits their ability to make decisions or to act to avoid injury. In the case of hidden contracts, firms that blur the traces of their previous contracts—without warning consumers or making this practice explicit—aggravate information asymmetries. Such information asymmetries undercut consumers' ability to make rationally informed decisions. In essence, we submit that there should be little question as to whether consumers can reasonably avoid the injury hidden contracts pose.

After establishing that consumers cannot reasonably avoid the harm of hidden contracts, the next step is to attend to the third and final prong of the unfairness analysis: whether the benefits to consumers or competition outweigh the injury. This component necessitates a cost-benefit analysis demonstrating that the practice at stake "is injurious in its net effects."[188]

When it comes to hidden contracts, it is hard to argue that such contracts benefit consumers or competition. Admittedly, firms might reduce operational costs by not maintaining an archive of previous contractual versions. However, these savings are unlikely to translate into lower prices for consumers, given the low cost of keeping such an archive. Firms already have these versions in an electronic format, and they all maintain websites with multiple pages. Thus, the costs of taking measures to prevent the injury—providing consumers with their original contracts and posting contractual versions online—should be minimal. In contrast, the harms hidden contracts inflict on consumers and society are considerable.[189]

Things are trickier, however, in terms of remedies. At the outset, it would be challenging for consumers and their advocates to quantify the damage and show how and to what extent hidden contracts prevented them from action and eroded their ability to seek remedies. At the same time, disgorgement would require evaluating the additional profits that hidden contracts generate. Things get even further complicated given that consumers did

of an obstacle to the free exercise of consumer decision making.' This definition was later adopted by courts and has become part of the accepted definition of the FTC's test for unfairness.") (quoting FED. TRADE COMM'N, *supra* note 175, ¶ 12).

188. *See* FED. TRADE COMM'N, *supra* note 175, ¶ 11.

189. *See* discussion *supra* Section I.B. (detailing the social costs of hidden contracts).

receive a product or a service from the firm that later hid its contract.

Therefore, one interesting path to consider is the idea of performance-based remedies.[190] According to this framework, a court could order a firm to comply with a "confusion injunction."[191] In essence, confusion injunctions ban "firms that have unfairly, deceptively, or abusively exploited customer confusion from continuing to do so."[192] In our context, a confusion injunction would require the firm to stop confusing consumers about their contractual rights and obligations by hiding the contract governing their business relationship. Likewise, a "consequences injunction" would "prohibit firms from continuing to unfairly, deceptively, or abusively inflict ill consequences on their customers."[193] Such an injunction would require the firm to not use hidden contracts to manipulate consumer decision-making.[194]

In summary, quantifying the injury that hidden contracts inflict on consumers and society and establishing causal links between hidden contracts and specific quantified harms may pose considerable challenges. These challenges are coupled with the weakness and gaps of UDAP laws and the political environment that may impact federal agencies and their appetite to act. In all, this analysis reinforces the need to propose a holistic approach. Such an approach should introduce a contract transparency duty, combining private, public, and administrative enforcement measures.[195]

IV. RESPONDING TO CRITIQUES

This Article proposes to impose a duty on firms to act transparently and provide a copy of their contracts to consumers. It also suggests that online firms should maintain a public archive with their historical agreements and that consumers should be

190. *See* Lauren E. Willis, *Performance-Based Remedies: Ordering Firms to Eradicate Their Own Fraud*, 80 LAW & CONTEMP. PROBS. 7, 8 (2017).

191. *Id.* at 8, 30.

192. *Id.* at 30.

193. *Id.*

194. *See id.*

195. *Cf.* Pridgen, *supra* note 165, at 933 ("Indeed, the use of private enforcement mechanisms as an extension of administrative agency regulation is a hallmark of the American regulatory system . . . ."); *Id.* at 935 ("The private right of action and the state powers of enforcement are truly complementary, not exclusive.").

entitled to receive a copy of their (original) contracts for a nominal fee. One may criticize this proposition from five major directions. The following sections present these criticisms and respond to them.

A Consumers' Personal Responsibility

First, one can argue that even if suppliers try to blur the tracks of their hidden contracts, consumers can (and perhaps should) save a copy of the agreement immediately after they accept it. Consumers should be responsible for their decisions and actions. A responsible consumer, the argument goes, should exercise vigilance and keep track of the contract she accepts. By keeping a copy of their contracts available, consumers actively care for their interests instead of relying on others, whether the firm or a government agency. When the need arises, consumers who save their original contract for future reference can retrieve it, learn about their rights and obligations, and make informed decisions.

As noted above,[196] we do not find this argument persuasive. Consumers typically agree to numerous standard form contracts.[197] Standard form contracts, especially online ones, govern multiple aspects of our everyday lives.[198] One accepts a standard form contract when engaging with others on social media, partaking in e-commerce, using dating apps, booking flights and hotels, purchasing insurance, opening a bank account, joining a gym, playing games online, or simply using public Wi-Fi—to name just a few. Moreover, consumer contracts are merely one type of legal document that people encounter. Other documents include, among other things, privacy policies, employment agreements, financial disclaimers, consent forms, and numerous disclosures. Can we reasonably expect consumers to track, sort, store, and be able to retrieve all these documents?

196. *See* discussion *supra* Section III.C (arguing that consumers cannot reasonably avoid the harm of hidden contracts by maintaining a personal inventory of form contracts).

197. *See* Alan Schwartz, *Proposals for Products Liability Reform: A Theoretical Synthesis*, 97 YALE L.J. 353, 372 n.36 (1988) ("[C]onsumers make many transactions."); Shmuel I. Becher & Tal Z. Zarsky, *Open Doors, Trap Doors, and the Law*, 74 L. & CONTEMP. PROBS. 63, 66 n.13 (2011) ("[C]onsumers engage in many transactions . . . .").

198. *See supra* note 1; *see also* Woodrow Hartzog, *Website Design as Contract*, 60 AM. U. L. REV. 1635, 1641 (2011) ("As websites became ubiquitous, so did terms of use. As a result, an overwhelming amount of online activity is not governed by default law but rather through agreement between the parties.") (footnote omitted).

There are many other reasons to doubt the personal responsibility proposition. For instance, consumers are typically time constrained,[199] and they often encounter form contracts when in a hurry or while being engaged with everything else life throws at us. Additionally, many consumers accept form contracts on their phones, where saving these documents can be less convenient and is not always intuitive or easy.

Cost-benefit analysis, cognitive biases, and limited mental bandwidth may further reduce consumers' tendency to diligently save consumer contracts. For example, each individual consumer may rationally believe that the probability that he or she will be a victim of an inefficient contract breach is low.[200] Hence, each individual consumer may rationally decide not to produce a digital or hard copy of each of the many contracts he or she accepts. Consumers may also believe that disagreements with sellers will be resolved in a friendly manner (or at least through simple correspondence, if not friendly), without the need to follow the strict arrangements of form contracts. Consumers may further assume that if worse comes to worst, they will be able to find the contracts they accepted online. At the time of contracting, consumers—as laypeople—may not envisage hidden contracts that follow unilateral amendments. People naturally tend to focus on the present,[201] discount future risks,[202] and be optimistic about the future.[203] There is no reason to penalize the average consumer and

199. Laura R. Bradford, *Emotion, Dilution, and the Trademark Consumer*, 23 BERKELEY TECH. L.J. 1227, 1283 (2008) ("Constraints on consumer's time . . . will always favor familiar brands because they will always seem less risky.") (footnote omitted). *Cf.* David Adam Friedman, *"Dishonest Search Disruption": Taking Deceptive-Pricing Tactics Seriously*, 51 U.C. DAVIS L. REV. ONLINE 121, 123 (2018) ("Consumers have time constraints . . . .").

200. *See* Kenneth K. Ching, *What We Consent to When We Consent to Form Contracts: Market Price*, 84 UMKC L. REV. 1, 3 (2015) ("Given the low probability that a dispute will arise over one of the unread terms . . . it would be 'irrational for form-receiving parties to spend time reading . . . the terms in the forms they sign.'") (quoting Randy E. Barnett, *Consenting to Form Contracts*, 71 FORDHAM L. REV. 672, 631 (2002)).

201. *See generally, e.g.*, Ted O'Donoghue & Matthew Rabin, *Doing it Now or Later*, 89 AM. ECON. REV. 103 (1999) (discussing the present bias); David Laibson, *Golden Eggs and Hyperbolic Discounting*, 112 Q. J. ECON. 443 (1997) (explaining how myopia can lead people to disproportionally care more about the present and not care enough about the future).

202. *See, e.g.*, Neil Weinstein, *Optimistic Biases About Personal Risks,* 246 SCIENCE 1232, 1232 (1989).

203. *See, e.g.*, Lynn A. Baker & Robert E. Emery, *When Every Relationship Is Above Average: Perceptions and Expectations of Divorce at the Time of Marriage*, 17 L. & HUM. BEHAV.

force them to be suspicious and litigious in every interaction with well-known online firms.

Finally, it is important to note how hidden contracts also undercut the potential of advanced technological tools ("smart readers") to assist consumers who wish to understand their standard form contracts.[204] Indeed, smart readers may be able to simplify contracts for consumers, personalize the content of these contracts, construct, and potentially benchmark them.[205] However, such smart readers need input (a contract text) to analyze. Hidden contracts entail no text at the outset, making even the most powerful and advanced technological language processing models useless.

B. Social Norms are Superior to Legal Rules

A second critical argument against our suggestion to introduce a transparency duty on suppliers may be that the duty is superfluous. According to this critique, consumers who need a copy of the original pre-amended contract can request it from the supplier. Once the consumer requires a copy of the agreement, the argument goes, the supplier will voluntarily agree to this request as an act of goodwill and to maintain its reputation. This will ease the burden on the supplier and ensure it will reproduce a copy of the original contract only for those consumers who actually want it. It will also eliminate regulatory and enforcement costs.

This line of reasoning is speculative and unsounded. First, it contradicts the impression one gets from online customers' complaints, which frequently protest that firms do not provide previous contractual versions.[206] Second, a consumer's email with an informal request may get lost in the firm's administrative maze, land in a spam or junk folder, or be inadvertently blocked or filtered. Such misfortune is less likely to occur when businesses are

439 (1993); Neil D. Weinstein & William M. Klein, *Unrealistic Optimism: Present and Future*, 15 J. SOC. & CLINICAL PSYCH. 1 (1996).

204. *See* Yonathan A. Arbel & Shmuel I. Becher, *Contracts in the Age of Smart Readers,* 90 GEO. WASH. L. REV. 83 (2022) (explaining how "Smart Readers" that employ language processing models such as GPT-3 can assist consumers and facilitate market competition over consumer form contract terms).

205. *Id.* at 94–109 (detailing the capabilities of smart readers).

206. *See, e.g.*, Live Better, *supra* note 33 ("We emailed seven different Epidemic Sound employees asking them for these older versions of their legal documents [including the terms of service], but none of them replied.").

formally obliged to provide consumers with a copy of their original contracts, because firms will establish filters or frameworks to address such inquiries. Third, and perhaps most importantly, as rational profit maximizers, companies often have weak incentives to voluntarily fulfill a consumer's request to receive a copy of the original contract.[207] Firms may be concerned that fulfilling the consumer's request may increase the probability that the consumer will raise demands, file third-party complaints, or take legal actions against the firm, based on the contractual legal terms. After all, why else would consumers ask for their original contracts? Hence, businesses may often prefer to keep consumers under the legal uncertainty that hidden contracts facilitate unless governed by a transparency duty. Without a clear rule mandating transparency, one should assume that firms would employ transparency strategically: opting for transparency only when it serves their interests (e.g., when the original form does not support the consumer's claim).

Furthermore, placing the burden on consumers to request the contract and subjecting them to the firm's will is problematic in and of itself.[208] As noted, most consumers lack legal knowledge, are generally unaware of their rights, are undermotivated to insist on those rights, and may fear confronting or suing the firm.[209] To increase access to justice, we should make it as easy and simple as

207. *Cf.* Eric J. Gouvin, *Truth in Savings and the Failure of Legislative Methodology*, 62 U. CIN. L. REV. 1281, 1340 (1994) ("[B]anks are profit-oriented enterprises and will not voluntarily give up an advantageous position . . . ."); Sarah L. Stafford, *Outsourcing Enforcement: Principles to Guide Self-Policing Regimes*, 32 CARDOZO L. REV. 2293, 2298 (2011) ("[P]rofit-maximizing facilities will not voluntarily undertake costly self-audits and turn themselves in for discovered violations unless there is some type of inducement for doing so.").

208. For a recent accessible account illustrating the hurdles that private enforcement of consumer protections law can entail see Jeff Guo, Alexi Horowitz-Ghazi, Willa Rubin & Keith Romer, *Spam Call Bounty Hunter*, NPR (Dec. 14, 2022, 6:43 PM), https://www.npr.org/2022/12/07/1141358550/spam-call-bounty-hunter-telemarketing (detailing the story of an individual who sought to file suits against firms that did not comply with the Do Not Call register) (last visited Oct. 3, 2023).

209. *Cf.* William L.F. Felstiner, Richard L. Abel & Austin Sarat, *The Emergence and Transformation of Disputes: Naming, Blaming, Claiming . . .* 15 L. & SOC'Y REV. 631, 651–52 (1980–1981). Consumers may remain passive and not partake in legal proceedings even as defendants. *See* Emily S. Taylor Poppe, *Why Consumer Defendants Lump It*, 14 NW. J. L. & SOC. POL'Y 149, 151 (2019) ("In the majority of actions to collect on consumer debts, default judgments are entered against consumer borrowers who fail to appear, even when there are flaws in the creditors' claims.").

possible for laypeople to access their legal documents and become acquainted with their rights.

If anything, providing firms with discretion about who receives a copy of the original agreement and under what circumstances may aggravate the undesirable distributional effects of hidden contracts. Consumers are heterogeneous. Consumers who vocally demand a contract and threaten a firm's reputation are likely to be privileged consumers (e.g., white, educated, male). At the same time, disadvantaged consumers are not likely to be assertive and insist on their rights.[210] This aggravates the negative impact of hidden contracts on marginalized consumers, further disempowering them and limiting their access to justice.

C. Original Contracts are Already Available Online

The third critique against our proposal is that consumers may locate their original contracts themselves via third-party online archives. If the original contracts are already available online, imposing a transparency duty on firms is superfluous. In fact, according to this line of reasoning, a transparency duty can backfire and harm consumers by increasing the costs of doing business. Firms would likely pass onto consumers the additional administrative costs that this duty generates.

We find this argument normatively inappropriate, factually inaccurate, and partially misleading. Normatively, much of the preceding analysis regarding access to justice and distribution effects is relevant here too. There is no basis to believe the average consumer has the knowledge and initiative to search for old form contracts via third-party archives. Moreover, placing this burden on individuals ignores the digital divide and harms the most vulnerable consumers.[211] Many consumers, including the elderly,

210. *See supra* note 35.

211. *See, e.g.*, Mark Lloyd, *The Digital Divide and Equal Access to Justice*, 24 HASTINGS COMMC'NS & ENT. L.J. 505, 527–30 (2002) (arguing that unequal access to technology could exacerbate inequalities). *See also* Emily A. Vogels, *Digital Divide Persists Even as Americans with Lower Incomes Make Gains in Tech Adoption*, PEW RSCH. CTR., (June 22, 2021), https://www.pewresearch.org/fact-tank/2019/05/07/digital-divide-persists-even-as-lower-income-americans-make-gains-in-tech-adoption (last visited Oct. 3, 2023); Anique Scheerder, Alexander van Deursen & Jan van Dijk, *Determinants of Internet Skills, Uses and Outcomes. A Systematic Review of the Second- and Third-Level Digital Divide*, 34 TELEMATICS & INFORMATICS 1607 (2017).

poor, and disadvantaged, may not be technologically savvy.[212] These consumers may be unaware of the advanced online tools that provide access to old versions of a website's contract terms.

But beyond that, the mere assumption that online archives provide a satisfactory source for previous versions of consumer contracts is factually incorrect. To be sure, some third-party online archives, such as "Wayback Machine,"[213] allow consumers to locate *some* old web pages that may include the original pre-amendment versions of contracts. However, these websites are an incomplete solution.

To begin with, online archives do not track the history of every webpage.[214] This incompleteness stems from two major reasons. First, the online contracting environment and the ways in which businesses and consumers interact online are constantly and rapidly evolving.[215] Second, website owners are allowed to exclude their sites from online archives.[216] Moreover, even when these archives include previous versions of a webpage, their historical tracking of a webpage is sometimes partial. To illustrate, while

212. *Cf.* Jennifer Lynch, *Identity Theft in Cyberspace: Crime Control Methods and Their Effectiveness in Combating Phishing Attacks*, 20 BERKELEY TECH. L.J. 259, 284 (2005) ("[B]ehavioral studies have shown that consumers are not instituting practices recommended by the FTC to protect themselves from online fraud, possibly because they are not technically savvy.") (citing George R. Milne, Andrew Rohm & Shalini Bahl, *Consumers' Protection of Online Privacy and Identity*, 38 J. CONSUMER AFFS. 217, 223–24 (2004)); (Jerry) Jie Hua, *Toward A More Balanced Model: The Revision of Anti-Circumvention Rules*, 60 J. COPYRIGHT SOC'Y U.S.A. 327, 348 (2013) ("Technologically-unsavvy consumers often are not conscious" of monitoring and data collection techniques used by firms.); John Shaeffer & Charlie Nelson Keever, *Privacy as a Collective Norm*, 41 LOY. L.A. ENT. L. REV. 253, 293 (2021) ("Big [D]ata is likely to affect the welfare of unsophisticated, vulnerable, and technologically unsavvy consumers more negatively.") (quoting Nir Kshetri, *Big Data's Impact on Privacy, Security and Consumer Welfare*, 38 TELECOMMS. POL'Y 1134, 1152 (2014) (alteration in original)).

213. WAYBACK MACHINE, https://archive.org/web (last visited Oct. 3, 2023).

214. TODD G. SHIPLEY & ART BOWKER, INVESTIGATING INTERNET CRIMES: AN INTRODUCTION TO SOLVING CRIMES IN CYBERSPACE 301 (2014) (Wayback Machine "does not crawl and record everything found on a website or webpage."); Ludovica Price, *Internet Archiving – The Wayback Machine*, HUMANITIES COMMONS 3 (2011) ("[T]he Archive cannot be considered to be 'complete.'"). As an example, Wayback Machine has not stored the URL of the contract terms of AliExpress.com, available at https://terms.alicdn.com/legal-agreement/terms/suit_bu1_aliexpress/suit_bu1_aliexpress202204182115_66077.html?spm=a2g0o.home.0.0.1ea66b05vMwOJG.

215. Price, *supra* note 214, at 3 ("But the size of the internet, its constant growth and mutability has since made net-wide crawls virtually impossible.").

216. *Wayback Machine General Information*, INTERNET ARCHIVE, https://help.archive.org/help/wayback-machine-general-information (last visited Oct. 3, 2023) ("Pages may not be archived due to robot's exclusions and some sites are excluded by direct site owner request.").

Yahoo.com was founded in 1994,[217] the tracking of its terms of use webpage only goes back to September 2021.[218] Beyond this, the archives do not store traditional offline consumer contracts, such as agreements between consumers and brick-and-mortar retailers (which may still employ digital means to form a contract; for example, a rental company that asks the consumer to sign digitally a contract presented on a screen).

Finally, it is also somewhat misleading to suggest that contracts found via online archives, such as "Wayback Machine," are fully evidentiarily acceptable. Courts are still grappling with the legal status of old versions of webpages found via online archives.[219] Though some courts find these sources acceptable, others do not. Given this issue is still in flux, consumers should not bear the risk that courts will reject the contract they present as evidentiarily deficient or inadequate.

D. Hidden or Not, Consumers Will Not Read

Another possible critique against our proposed recommendations is that whether form contracts are accessible or not, consumers do

217. *Yahoo!*, BRITANNICA, https://www.britannica.com/topic/Yahoo-Inc (last visited Oct. 3, 2023).

218. *See* the URL of Yahoo's terms of use, https://legal.yahoo.com/us/en/yahoo/terms/otos/index.html (last visited Oct. 3, 2023). According to Wayback Machine, it started saving this URL on September 1, 2021. *See* https://web.archive.org/web/20220000000000*/https://legal.yahoo.com/us/en/yahoo/terms/otos/index.html (last visited Oct. 3, 2023).

219. *See, e.g.*, Ward v. Am. Airlines, Inc., No. 4:20-CV-00371-O, 2020 WL 8300505, at *1–2 (N.D. Tex. Oct. 16, 2020) ("[T]he Court declines to take judicial notice of Hotwire's and Expedia's Terms of Use by way of the Internet Archive Wayback Machine."); Lloyd v. The Retail Equation, Inc., No. CV 21-17057, 2022 WL 18024204, at *9 (D.N.J. Dec. 29, 2022) ("[I]t is unclear whether screenshots from the Wayback Machine are admissible in the Third Circuit absent authentication by an Internet Archive employee."); Weinhoffer v. Davie Shoring, Inc., 23 F.4th 579, 584 (5th Cir. 2022) ("[O]ther district courts have held that evidence from the Wayback Machine 'is not so reliable and self-explanatory that it may be an appropriate candidate for judicial notice . . . .'") (quoting My Health Inc. v. Gen. Elec. Co., No. 15-CV-80-JDP, 2015 WL 9474293 at *4 (W.D. Wis. Dec. 28, 2015). *But see* Thorne v. Square, Inc., No. 20-CV-5119(NGG)(TAM), 2022 WL 542383, at *1 (E.D.N.Y. Feb. 23, 2022), appeal withdrawn, No. 22-542, 2022 WL 2068771 (2d Cir. Apr. 14, 2022) ("The court therefore joins the chorus of others and takes judicial notice of these archived webpages from the Wayback Machine as the relevant terms of service that Plaintiffs would have respectively viewed when registering for Cash App and requesting a Cash Card."); Gardiner v. Walmart Inc., No. 20-CV-04618-JSW, 2021 WL 2520103, at *9 (N.D. Cal. Mar. 5, 2021) ("Courts have routinely taken judicial notice of contents of web pages available through the Wayback Machine as facts that can be accurately and readily determined from sources whose accuracy cannot be questioned.").

not, and will not, read them. Undeniably, evidence suggests that ordinary consumers, law professors, and judges all do not read consumer contracts.[220] If no one reads form contracts, the argument goes, there is little wisdom in endeavoring to make them available and accessible. If anything, these efforts are doomed to fail and are thus wasteful.

With respect, this argument misses the mark. First and foremost, as the case of the Palmers illustrates, our analysis focuses on the ex-post stage—once consumers experience a problem with the product, service, or firm and need to access their (hidden) contracts. While consumers may not read their contracts ex ante, they reveal a stronger tendency to read and act upon their contracts once a dispute or a problem arises.[221] Indeed, consumers often share their inability to find their form contracts.[222] Hence, making contracts available (i.e., unhiding contracts) can serve consumers ex post, regardless of their tendency to ignore form contracts ex ante.

Second, we noted in section IV.A the potential of new technologies ("smart readers") to assist consumers with their form contracts.[223] The possibility of using such smart readers entails freeing consumers from the need to read long and complex arrangements. Instead, consumers can delegate such tasks to machines. However, if a contract is hidden, smart readers do not

220. *See, e.g.*, Bakos et al., *supra* note 7 (finding that virtually all online users do not read EULAs); Schmitz, *supra* note 8, at 873–78 (reviewing empirical research suggesting consumers do not read form contracts); Jeff Sovern, *The Content of Consumer Law Classes III*, 22 J. CONSUMER & COM. L. 2, 4 (presenting survey results indicating that 57% of consumer law professors "rarely or never" read consumer contracts); Debra Cassens Weiss, *Chief Justice Roberts Admits He Doesn't Read the Computer Fine Print*, A.B.A. J. (Oct. 20, 2010), https://perma.cc/964P-EGVW (last visited Oct. 3, 2023).

221. *See, e.g.*, Shmuel I. Becher & Tal Z. Zarsky, *E-Contract Doctrine 2.0: Standard Form Contracting in the Age of Online User Participation*, 14 MICH. TELECOMM. & TECH. L. REV. 303, 315 (2008) ("Most of the reasons for the lack of effective reading and comprehension of 'non-salient' terms ex ante do not apply to the ex post context."); Shmuel I. Becher & Esther Unger-Aviram, *The Law of Standard Form Contracts: Misguided Intuitions and Suggestions for Reconstruction*, 8 DEPAUL BUS. & COM. L.J. 199, 214–15 (2010) (Many more consumers indicated "they would read the contract ex post (rather than ex ante) . . . ."); Meirav Furth-Matzkin, *On the Unexpected Use of Unenforceable Contract Terms: Evidence from the Residential Rental Market*, 9 J. LEG. ANALYSIS 1, 7 (2017) (finding that tenants are likely to be deterred by the terms of their lease agreements once a dispute arises even if those terms are unenforceable).

222. *See, e.g.*, examples noted *supra* note 33 (detailing examples of users who cannot find their contracts); *see also* cases cited *supra* note 215 (illustrating the need to sometimes use the Wayback Machine to present the governing contract to the court).

223. *See* discussion *supra* text accompanying notes 200–201.

have the necessary input to analyze it and thus cannot help consumers.[224] Therefore, hidden contracts subvert consumers' ability to use new technologies to effectively cope with standard form contracts.

Third, throughout this Article we explained how hidden contracts harm not only consumers but also consumer organizations, intermediaries, and other legal and extralegal forces. Therefore, focusing only on consumer readership (or lack thereof) neglects other important pieces of the puzzle. Overall, lowering the search costs for those consumers and third parties who wish to access consumer form contracts is a legitimate goal in and of itself.

E. The Original Contract Is Irrelevant

Finally, a seemingly potent yet factually inaccurate criticism opines that we should not care too much about the original contract the consumer accepted. As the argument goes, once suppliers modify the contract, the original contract becomes legally irrelevant. According to this logic, the new modified contract invalidates and replaces the original one. It is thus futile to invest resources to allow consumers access to the original contract.

This critique has some merit, but it is not entirely persuasive. To be sure, U.S. law allows firms significant discretion to amend their contracts.[225] However, as the case of the Palmers suggests,[226] unilateral modification of the contract does not automatically nullify and substitute the original one. Typically, contractual amendments should satisfy various legal requirements in order to replace the original agreement.

Notably, the recipient of the proposed contractual changes must receive a reasonable (1) notice of the proposed modified terms, (2) opportunity to review these terms, and (3) opportunity to reject the proposed modified terms (or terminate the agreement

224. *See* Arbel & Becher, *supra* note 200 (explaining the power and promise of GPT-3 smart readers).

225. *See* RESTATEMENT (SECOND) OF CONTRACTS § 89 (AM. L. INST. 1981) (detailing the circumstances and allowed scope of contractual modifications); U.C.C. § 2-209 (AM. L. INST. & NAT'L CONF. COMM'RS ON UNIF. STATE L. 2020) (stating that modification needs no consideration and noting basic requirements and limitations).

226. *See* discussion *supra* text accompanying notes 19–30.

with no penalty).[227] In addition, for a modification to replace the original contract, a consumer must either manifest assent to the modified terms or not reject them.[228] Moreover, the modifying party must propose the contractual amendment in good faith, limiting firms to modifications within the parties' original reasonable expectations.[229] Likewise, modifications shall not undermine an affirmation or promise made by the business that was a central part of the original bargain between the contracting parties.[230] Courts apply these criteria quite consistently and often reject contractual modifications. [231]

To be sure, courts and commentators vary in the importance they attribute to notice and the degree to which notice should affect enforceability. Judicial decisions regarding modifications also depend on the concrete circumstances of the individual case, which courts examine on a case-by-case basis. That said, courts tend to ensure that firms properly communicate the changes to consumers and that consumers have a reasonable opportunity to reject the modifications.[232] For example, merely posting modified terms online does not constitute reasonable notice.[233] Likewise, a message buried in fine print and hidden at the bottom of dense text does not

227. RESTATEMENT OF THE LAW OF CONSUMER CONTRACTS 66 (AM. L. INST., Tentative Draft No. 2, 2022).

228. *Id.*

229. *See e.g.*, Badie v. Bank of America, 79 Cal. Rptr. 2d 273, 284 (Cal. Ct. App. 1998) ("[I]t is the Bank's exercise of its discretionary right to change the agreement . . . which must first be analyzed in terms of the implied covenant [of good faith and fair dealing]."); Cycle City, Ltd. v. Harley-Davidson Motor Co., 81 F. Supp. 3d 993, 1014 (D. Haw. 2014) ("Harley-Davidson reserved the right to change prices, but, in doing so, it was obligated to act in good faith. When an express contract provision allows a party to exercise some discretion, that party is obligated to exercise its discretion in good faith.") (citing Damabeh v. 7-Eleven, Inc., No. 12-CV-1739-LHK, 2013 WL 1915867, at *6 n.4 (N.D. Cal. 2013).

230. RESTATEMENT OF THE LAW OF CONSUMER CONTRACTS *supra* note 227.

231. In the sample of cases surveyed by the Restatement of the Law of Consumer Contracts, courts rejected modifications in almost 40 percent (38 of 97) of the cases.

232. RESTATEMENT OF THE LAW OF CONSUMER CONTRACTS *supra* note 227.

233. *See e.g.*, Rodman v. Safeway Inc., No. 11-cv-03003-JST, 2015 U.S. Dist. LEXIS 17523, at *10 (N.D. Cal. Feb. 12, 2015) ("Even if a customer's continued use of a service could be considered assent to revised terms, 'such assent can only be inferred after [that customer] received proper notice of the proposed changes.'") (quoting Douglas v. U.S. Dist. Court for Cent. Dist. Of Cal., 495 F.3d 1062, 1065 (9th Cir. 2007) (alteration in original); *Douglas*, 495 F.3d at 1066 ("Even if Douglas's continued use of Talk America's service could be considered assent, such assent can only be inferred after he received proper notice of the proposed changes.").

suffice.[234] Similarly, posting the changes on a remote webpage or including them in an inconspicuous hyperlink obscured with many other links would not satisfy the requirement for proper notice.[235]

Furthermore, even if the modification satisfies these requirements and replaces the original contract, there is still considerable value in accessing the original contract and tracking the changes the firm made. Watchdogs, consumer organizations, the media, policymakers, regulators, and academics may all wish to access and investigate previous versions of consumer form contracts. Contractual archives would contribute valuable intelligence for monitoring and supervision agencies, which may indicate which practices or industries merit vigilance and prioritizing.[236] All in all, exploring contractual modifications can highlight or expose important trends and trajectories, which studying specific (current) contracts in isolation cannot reveal. Such trends can pertain, for example, to how firms structure their dispute resolution processes (mandated arbitration, jury waivers, class action limitations, forum selection clauses); the changes in length and complexity of contracts throughout the year; the ways firms respond to case law, and much more.

To sum up, there are legitimate justifications to ensure access to previous consumer form contract versions. First and foremost, earlier versions may be legally binding and relevant where a contract modification does not fulfill all the legal requirements. Moreover, consumers may wish to compare revised terms with old ones in deciding whether to continue their relationships with the

234. *See e.g.*, Murray v. Grocery Delivery E-Services USA Inc., 460 F. Supp. 3d 93, 98 (D. Mass. 2020) ("[T]he supposed notice was not notice at all . . . .[It] was given by a one-line statement at the bottom of the email . . . ."). *See also* Martin v. Comcast, 146 P.3d 380, 389 (Or. Ct. App. 2006) (rejecting a modification included in bill stuffers since consumers were not provided with sufficient notice); Powertel, Inc. v. Bexley, 743 So. 2d 570, 575 (Fla. Dist. Ct. App. 1999) (rejecting a modification sent as an insert with the customer's bill, finding it unconscionable).

235. *See e.g.*, Grosvenor v. Qwest Corp., 854 F. Supp. 2d 36 1021, 1034 (D. Colo. 2012) ("[T]he Court is not convinced that a simple requirement that Qwest post any changes it makes to its agreement to a remote webpage is material."); In re Zappos.com, Inc., Customer Data Sec. Breach Litig., 893 F. Supp. 2d 1058, 1066 (D. Nev. 2012) ("A party cannot assent to terms of which it has no knowledge or constructive notice, and a highly inconspicuous hyperlink buried among a sea of links does not provide such notice.").

236. *See also supra* note 139 (noting the CFPB's ability to use data from a registry of contract terms to identify risks that terms and conditions pose to consumers and prioritize and scope regulatory measures).

firm or comprehend new risks, rights, and obligations. Finally, previous contractual versions have value beyond the particular case, serving diverse stakeholders that monitor, scrutinize, and report on possible trajectories and trends, and discipline firms' behavior.

CONCLUSION

Giant websites often declare that they are committed to the principle of transparency.[237] This Article, however, reveals that such websites too often utilize the benefits of online contracting yet apply nontransparent contractual practices. These practices obfuscate the contents of their standard form agreements and facilitate what this Article dubs "hidden contracts."[238]

Hidden contracts undermine access to justice while harming the most vulnerable consumers. Such contracts reduce consumers' ability to assess the contracting parties' rights and obligations and make informed decisions about potential disputes and grievances.[239] Hidden contracts prevent consumers from making informed decisions about insisting on their rights, complaining about firms, sharing their experiences online, and bringing their cases to court. This uncertainty may increase the suppliers' incentive to breach their contracts at the expense of consumers.[240] Furthermore, hidden contracts also impede the ability of other interested parties to study consumer contracts, identify concerning trends, monitor firms, and address possible regulatory loopholes.[241]

This Article focused on the contractual practices of 100 of the most popular websites. These websites are rather heterogenous and govern practical aspects in the everyday lives of billions of consumers.[242] Given the importance of these websites and how the online sphere rules much of our lives, future research could examine upscaling and broadening our analysis. We believe that much of our examination may apply to additional documents, including precontractual promises, risk disclosures, consent forms,

237. *See* discussion *supra* notes 12–14 and accompanying text.

238. *See* discussion *supra* Section II.C.

239. *See* discussion *supra* Section I.B.

240. *See id.*

241. *See* discussion *supra* Section IV.D (noting the potential interest of multiple stakeholders in previous version of consumer stand form contracts).

242. *See* discussion *supra* Section II.A.

privacy agreements, behavioral policies, and codes of conduct. At the same time, we also acknowledge the limitations of our sample. There are many types of consumer form contracts, including offline contracts, and our sample may not capture them all. We thus hope that future research will shed further light on the prevalence, danger, implications, mechanics, and regulation of hidden contracts.

As a first step, and bearing in mind the ways technology eases and facilitates information sharing and retention, we suggest introducing a contract transparency duty.[243] According to this proposal, firms would (1) supply copies of contracts to consumers shortly after consumers accepts them; (2) maintain previous versions of their consumer form contracts on their websites; and (3) provide consumers with copies of the original contracts they accepted whenever consumers request them, charging merely nominal fees.

Admittedly, a transparency duty cannot guarantee a fair overall market equilibrium. Nonetheless, it would mark a considerable step toward empowering consumers, improving access to justice, materializing the untapped potential of transparency, and more justly distributing the benefits of the online environment.

243. *See* discussion *supra* Section III.

APPENDIX A: SAMPLE WEBSITES (ALPHABETICALLY LISTED BY COLUMN)

1. adobe.com	35. genius.com	69. sears.com
2. alamy.com	36. github.com	70. shutterstock.com
3. alibaba.com	37. glassdoor.com	71. simplyhired.com
4. aliexpress.com	38. goodhousekeeping.com	72. soundcloud.com
5. amazon.com	39. goodreads.com	73. spotify.com
6. bbc.com	40. google.com	74. springer.com
7. bestbuy.com	41. homedepot.com	75. stackexchange.com
8. bloomberg.com	42. houzz.com	76. stackoverflow.com
9. britannica.com	43. imdb.com	77. tandfonline.com
10. businessinsider.com	44. indeed.com	78. target.com
11. buzzfeed.com	45. insider.com	79. thefreedictionary.com
12. buzzfile.com	46. instagram.com	80. theguardian.com
13. cambridge.org	47. issuu.com	81. tiktok.com
14. cbsnews.com	48. istockphoto.com	82. time.com
15. chamberofcommerce.com	49. jstor.org	83. trip.com
16. chron.com	50. latimes.com	84. tripadvisor.com
17. cnbc.com	51. linkedin.com	85. tumblr.com
18. cnn.com	52. manta.com	86. twitter.com
19. costco.com	53. merriam-webster.com	87. usnews.com
20. cylex.us.com	54. microsoft.com	88. walmart.com
21. dailymail.co.uk	55. npr.org	89. washingtonpost.com
22. dailymotion.com	56. nypost.com	90. wayfair.com
23. deviantart.com	57. nytimes.com	91. webmd.com
24. dreamstime.com	58. oup.com	92. weebly.com
25. ebay.com	59. pinterest.com	93. wikihow.com
26. espn.com	60. popsugar.com	94. wordpress.com

27. etsy.com	61. quizlet.com	95. yahoo.com
28. expedia.com	62. quora.com	96. yellowpages.com
29. facebook.com	63. redbubble.com	97. yelp.com
30. fandom.com	64. reddit.com	98. youtube.com
31. flickr.com	65. redfin.com	99. zillow.com
32. forbes.com	66. researchgate.net	100. ziprecruiter.com
33. foursquare.com	67. reuters.com	
34. gamespot.com	68. sciencedirect.com	

Twenty-First Century Split: Partisan, Racial, and Gender Differences in Circuit Judges Following Earlier Opinions

Stuart Minor Benjamin, Kevin M. Quinn & ByungKoo Kim[*]

Judges shape the law with their votes and the reasoning in their opinions. An important element of the latter is which opinions they follow, and thus elevate, and which they cast doubt on, and thus diminish. Using a unique and comprehensive dataset containing the substantive Shepard's treatments of all circuit court published and unpublished majority opinions issued between 1974 and 2017, we examine the relationship between judges' substantive treatments of earlier appellate cases and their party, race, and gender. Are judges more likely to follow opinions written by colleagues of the same party, race, or gender? What we find is both surprising and nuanced. We have two major findings. First, over the forty-four-year span we studied, we find growing partisan differences in positive treatments of earlier cases. The partisan differences are largest for treatments in ideologically salient categories of cases. Interestingly, the partisan differences arise more for treatments of opinions written by Democratic

* Benjamin is the William Van Alstyne Professor of Law at Duke Law School; Quinn is a Professor at Emory Law School and in the Department of Quantitative Theory and Methods at the Emory College of Arts and Sciences; Kim is an Assistant Professor at the KDI School of Public Policy and Management. More than the usual thanks are in order. We thank LexisNexis, who shared their data with us in 2018 and patiently tracked down the answers to our many questions about the data over the next twenty-seven months. A dataset so massive required similarly massive work to identify, categorize, and organize its elements. Tom Balmat at Duke and Sean Chen at the Duke Law Library were particularly valuable in that lengthy process. Jane Bahnson and Wickliffe Shreve, also at the Duke Law Library, were indefatigable in helping us find relevant materials about *Shepard's*, Lexis, and court practices. Gloria Han, Andrew Tisinger, and Wenyi Zhou provided excellent research assistance, and Leanna Doty and Balfour Smith provided excellent editing assistance. Over the five-year gestation of the project, we have benefited from many colleagues' comments, including Matt Adler, Joseph Blocher, Curt Bradley, Michael Frakes, Jack Knight, Maggie Lemos, Marin Levy, Mat McCubbins, Darrell Miller, Jeff Powell, Neil Siegel, Georg Vanberg, Jonathan Wiener, and Ernie Young. We also thank colleagues who attended our 2021 Duke Law workshop on an early version of this Article as well as members of the Emory Law faculty who attended a 2022 workshop on a version of this Article.

appointees than for opinions written by Republican appointees, which we think is best explained by an accelerating movement among Republican appointees in a conservative direction compared to a steady move among Democratic appointees in a liberal direction. The increase in partisan differences is not a function of presidential cohorts or age cohorts. Second, there are intraparty racial and gender differences in positive treatments of past cases, and these differences are similar to the partisan differences. Within each party, Black and White judges differ in their treatments of opinions authored by Black co-partisans, Hispanic and White judges differ in their treatments of opinions authored by Hispanic co-partisans, and female and male judges differ in their treatments of opinions authored by female co-partisans. Similar to the partisan divergence noted above, we also find that some of these differences increase in magnitude over time—with particularly notable increases in the Black-White Democratic differences, Hispanic-White Republican differences, and female-male Republican differences. Notably, the racial and gender differences we find in positive Shepard's treatments are not mirrored in most studies of racial and gender differences in judicial behavior, which focus on merits votes and include a much smaller number of cases.

These results defy easy explanation. They do not support the proposition that party, race, and gender have always played a pervasive role for judges. Instead, our results provide evidence of increasing partisan, racial, and gender polarization among judges in recent years. For reasons we explain in the body of this Article, the partisan, racial, and gender differences we find appear to be a function of political ideology. Further, because the racial and gender differences are within parties, our results indicate that not only partisan differences but also intraparty racial and gender ideological differences have risen in recent years (particularly for Republican judges).

Our data thus reveal polarization among circuit judges and, as a result, in their shaping of the law. Many groups in the United States have become more ideologically polarized in recent years. Our data indicate that judges are one of them.

CONTENTS

INTRODUCTION

Many researchers have studied the relationship between judges' behavior and their political party, race, or gender. Most of these studies focus on whether judges' votes in a particular set of cases are associated with the judges' party, race, or gender.[1] As we discuss in more detail in Part II, these studies have generally found

1. Throughout this Article, we generally use the term "race" to refer to both race and ethnicity. And, as we discuss below, in determining party we look to the political party of the President who most recently appointed the judge. *See infra* note 42 and accompanying text. To avoid wordiness, we sometimes refer to Democratic and Republican appointees by the shorthand "Democrats" and "Republicans."

that Democratic and Republican appointees vote differently in some categories of ideologically salient cases, that judges of different races vote differently in a smaller subset of cases, and that male and female judges vote differently in an even smaller subset of cases.[2]

Judges' votes are very important, but judges of course issue opinions, and the reasoning in those opinions helps to shape later cases. One significant aspect of a given opinion is which earlier cases it follows and which it casts doubt on. Using an original dataset containing all the substantive *Shepard's Citations* (*Shepard's*) treatments in all federal appellate majority opinions from 1974–2017 (670,784 opinions in total), we examine how a judge's party, race, and gender are associated with changes in an opinion's likelihood of following an earlier circuit opinion.[3]

The vast amount of data we have allows us to make adjustments for each of the 521 circuit-year combinations in our data.[4] This contrasts with most empirical studies of federal appellate decision making, which do not make such adjustments and include far fewer cases (usually only hundreds of cases).[5] We make these adjustments for all our analyses to reduce the possibility of producing biased results (e.g., by one circuit having a larger number of judges of a particular party, race, or gender in earlier years and another circuit having a larger number in later years, or by turnover on party, race, or gender lines within a circuit over the long period of our study).

2. *See infra* notes 42–68 and accompanying text; on "male" and "female," *see infra* note 7.

3. We, and *Shepard's*, focus on majority opinions. *See infra* notes 28 and 99 and accompanying text. For conciseness, we use the term "opinions" to refer to majority opinions. And because we are focusing on majority opinions, we largely use the terms "opinion" and "case" synonymously.

4. There are 521 circuit-year combinations because our data cover forty-four years and the twelve regional circuits (we do not include the Federal Circuit because it has relatively few of the ideologically, race-, and gender-salient cases that we want to measure; *see infra* text accompanying note 100). The Eleventh Circuit did not exist in the first seven years of our data (because it was part of the Fifth Circuit), thus yielding 521 circuit-year combinations instead of 528.

5. *See infra* notes 67–68 and accompanying text.

What might one expect to find? At one extreme, we might expect no meaningful differences related to partisanship,[6] race,[7] or gender in how later judges substantively treat earlier opinions: judges will follow, say, the canonical case rejecting a claim of ineffective assistance of counsel, and that canonical case will not have elements more likely to appeal to a later judge of the same party, race, or gender as the opinion author. The idea is that when judges choose which opinions to follow, party, race, and gender are irrelevant. On this account, judges are not influenced by the party, race, or gender of the authors whose opinions they follow (and may not even notice the party, race, or gender of the earlier author), and nothing in the earlier opinions of a judge of a particular party, race, or gender will be correlated with anything that a later judge might value.[8] If this account is correct, we would not expect to see any differences in substantive treatments correlated with party, race, or gender.

At the other extreme, we might expect pervasive differences related to party, race, and gender in how later judges substantively treat earlier opinions. The idea is that judges can choose among different opinions on ineffective assistance of counsel (to stick with the example), and they will tend to follow opinions written by judges with whom they share a party, race, or gender, because

6. We use the term partisanship simply to refer to political parties, not in the more informal sense of particularly strong support for a party or cause.

7. Our analyses of racial differences focus on Hispanic, Black, and non-Hispanic White judges because of the small number of judges in our data who self-identify with other racial/ethnic groups. We refer to Hispanic rather than Latinx judges because we are using the Federal Judicial Center's definitions. Relatedly, we use the gender binary "female" and "male" in referring to judges because the Federal Judicial Center uses only those categories for gender and there are no known transgender federal circuit judges.

We follow the most common conventions in judicial behavior studies in referring to judges appointed by Republican (Democratic) Presidents as Republican (Democratic) "appointees," because some judges may not be members of the President's party. But when discussing race and gender we use the term "judges," because the Federal Judicial Center data rely on judges' self-identified race and gender and we have no reason to doubt that self-identification. *See infra* text accompanying note 96; CASS R. SUNSTEIN, DAVID SCHKADE, LISA M. ELLMAN & ANDRES SAWICKI, ARE JUDGES POLITICAL?: AN EMPIRICAL ANALYSIS OF THE FEDERAL JUDICIARY 5–14 (2006).

8. Note that this second proposition is entailed in this hypothesis of no meaningful differences. If, for example, a Republican appointee preferred to follow a substantively conservative opinion, then we might expect that Republican appointee to be more likely to follow an opinion written by another Republican even if the later judge was unaware of the earlier judge's party, on the theory that opinions written by Republicans are, on average, more likely to be substantively conservative than are opinions written by Democrats.

the language and reasoning in the opinions they follow are closer to the later judges' preferences. As with the possibility that there are no differences based on party, race, or gender, it need not be that judges pay attention to these characteristics of the judges whose opinions they follow. The later judge may simply find the reasoning of a particular earlier opinion attractive for reasons correlated with these characteristics and thus be more likely to choose to follow an opinion written by a judge of the same party, race, or gender even if the judge ignores the identity of the author of the earlier opinion. So a later judge's awareness of the party, race, or gender of an opinion author is not necessary to motivate this account. But for this account to be plausible without such awareness, there must be some element of an opinion correlated with party, race, or gender (such as ideology) such that a later judge is likely to prefer to follow an opinion written by a judge of the same party, race, or gender. If this account is correct, then we would expect to see pervasive differences in substantive treatments correlated with party, race, and gender: judges will be more likely to (consciously or unconsciously) follow opinions written by those of the same party, race, or gender.

Our findings are surprising—and more nuanced than either of these accounts would suggest. Pooling data over the full time span of our data (1974–2017), we find statistically significant, but substantively small, partisan differences and intraparty racial and gender differences in authoring judges' positive treatments of earlier opinions.[9] But looking at the pooled data masks a dramatic development that is the real story: Far from being stable over time, the differences we find are quite small in the early years of our study and rise dramatically over time, becoming large and thus substantively meaningful.[10] Further, the differences are greatest for

9. These findings are for opinions that follow an earlier opinion and thus treat that earlier opinion as controlling. *See infra* notes 30–35 and accompanying text. We do not find substantively meaningful or statistically significant differences in negative treatments of opinions, which likely reflects the smaller number of such negative treatments (especially given our controls for circuit and year). *See infra* text accompanying note 101.

10. Statistical significance refers to the ability to reject a particular null hypothesis (typically of no difference or no effect) regardless of the size of the difference or effect. Substantive significance refers to an estimate that is large enough to be of scientific or policy interest. With a large enough sample, minuscule differences or effects can be statistically significant, even if they are of no scientific or policy interest. *See* STEPHEN T. ZILIAK & DEIRDRE N. MCCLOSKEY, THE CULT OF STATISTICAL SIGNIFICANCE: HOW THE STANDARD ERROR COSTS US JOBS, JUSTICE, AND LIVES 31–32 (2008) (drawing this distinction); Richard Lempert, *The*

the most ideologically charged categories of cases. Put differently, we do not see the sort of pervasive partisan, racial, and gender differences over the full time period that would exist if judges were consistently influenced by these factors. Instead, we see a sharp rise in partisan differences, and we see a rise in racial and gender differences within parties (particularly for Republicans). These differences are most dramatic in cases with the most ideological salience. The fact that these racial and gender differences occur within parties highlights that these differences are not the result of statistical associations between partisanship and race or between partisanship and gender. To pick the clearest example, Black and White co-partisan judges treat opinions by Black co-partisans differently in ways that are not only statistically significant but also large and therefore substantively meaningful.

A closer look at our results reveals that the partisan differences in substantive treatments of opinions written by Democratic appointees are larger than the differences in treatments of opinions by Republican appointees. What could explain greater partisan differences in treatments of Democratic opinions than in treatments of Republican opinions? The best explanation involves an accelerating shift among Republican appointees in a conservative direction compared to a steady shift among Democratic appointees in a liberal direction. Such a pattern produces results strikingly similar to what our data show. To be clear, we cannot prove this explanation, but we think it is the most likely one.

Finally, the race and gender findings (which, again, are within party) are particularly interesting, because our study of substantive treatments finds significant Black-White differences and significant Hispanic-White and female-male differences among Republicans, whereas studies of voting have found relatively few significant differences across judges of different races and genders.[11] As with partisanship, these differences have risen in recent years and apply to our broad category of ideologically salient cases (not just the

Significance of Statistical Significance: Two Authors Restate an Incontrovertible Caution. Why a Book?, 34 L. & SOC. INQUIRY 225, 227 n.4 (2009).

We do not use the terms "substantive significance" or "substantively significant" in this Article. The reason is to avoid confusion with the similar sounding but distinct "statistical significance" and "statistically significant." In place of "substantively significant" we use "substantively meaningful."

11. *See infra* notes 51–66 and accompanying text.

subset of ideologically salient cases that are race-salient and gender-salient).

We think the best explanation of these racial and gender differences is that they are capturing an element of ideology that partisanship does not capture. Our results are not consistent with a desire of judges of a particular race or gender to enhance the status of those of the same race or gender, because that would not explain the rise over time. The temporally increasing racial and gender differences in our data are similar to the increasing partisan differences we find. Both increases align with the widely documented rise in polarization among U.S. elected officials and within U.S. society more broadly.[12] There is evidence that just as a partisan affiliation reveals information about political attitudes and ideology,[13] race and gender are associated with political attitudes and ideology.[14] In light of these correlations, intraparty race and gender provide a finer-grained proxy for a judge's ideological leanings. Consequently, the fact that we see Hispanic-White and female-male differences only within Republican judges provides some support for the proposition that White male Republican appointees are the central contributors to the accelerating rightward move among Republican appointees that we find.[15]

12. *See* Alan I. Abramowitz & Kyle L. Saunders, *Is Polarization a Myth?*, 70 J. POL. 542, 546–47 (2008); ALAN I. ABRAMOWITZ, *Partisan-Ideological Polarization*, *in* THE DISAPPEARING CENTER: ENGAGED CITIZENS, POLARIZATION, AND AMERICAN DEMOCRACY 34, 37, 39–40 (2010); Drew DeSilver, *The Polarization in Today's Congress Has Roots that Go Back Decades*, PEW RSCH. CTR. (Mar. 10, 2022), https://www.pewresearch.org/fact-tank/2022/03/10/the-polarization-in-todays-congress-has-roots-that-go-back-decades; MICHAEL DIMOCK, JOCELYN KILEY, SCOTT KEETER & CARROLL DOHERTY, POLITICAL POLARIZATION IN THE AMERICAN PUBLIC: HOW INCREASING IDEOLOGICAL UNIFORMITY AND PARTISAN ANTIPATHY AFFECT POLITICS, COMPROMISE AND EVERYDAY LIFE 19–20, 24 (2014), https://www.pewresearch.org/politics/2014/06/12/political-polarization-in-the-american-public; Cynthia R. Farina, *Congressional Polarization: Terminal Constitutional Dysfunction?*, 115 COLUM. L. REV. 1689, 1697 (2015); Jeff Lewis, *Polarization in Congress*, VOTEVIEW.COM (Jan. 20, 2022), https://voteview.com/articles/party_polarization; *see also* Jennifer McCoy & Benjamin Press, *What Happens When Democracies Become Perniciously Polarized?*, CARNEGIE ENDOWMENT INT'L PEACE (Jan. 18, 2022), https://carnegieendowment.org/2022/01/18/what-happens-when-democracies-become-perniciously-polarized-pub-86190.

13. *See, e.g.*, DIMOCK ET AL., *supra* note 12 at 16.

14. *See, e.g.*, Vincent L. Hutchings & Nicholas A. Valentino, *The Centrality of Race in American Politics*, 7 ANN. REV. POL. SCI. 383, 401 (2004); *Gender Gap Public Opinion*, CTR. FOR AM. WOMEN & POL., https://cawp.rutgers.edu/gender-gap-public-opinion (last visited Oct. 3, 2023).

15. This would comport with some suggestive evidence regarding the preconfirmation political ideology of district court judges (as opposed to their judicial behavior) indicating

Part of what is striking about the racial and gender differences in substantive treatments is that prior research on merits votes has found differences in voting behavior between female and male judges, or judges of different races, only in limited subsets of cases. By contrast, the intraparty race and gender differences we find suggest the existence of subtle ideological differences operating outside of race- or gender-salient cases—differences that previous studies have failed to identify.

Our results do not support the proposition that judges have always been pervasively influenced by party, race, or gender. Instead, our results provide evidence of increasing partisan, racial, and gender differences among judges in recent years that reflect some combination of greater ideological differences and a greater willingness to let ideological differences influence how opinions are written.[16] Judges have some insulation from the increasing ideological polarization in the country, but that insulation goes only so far.

The rest of the Article proceeds as follows. In Part I, we discuss the importance of majority opinions' substantive treatments of earlier opinions. Substantive treatments are not mere citations. Following an opinion means relying on it as controlling authority. *Shepard's* is the most studied and accepted source of substantive treatments, and we rely on it here. Part II discusses the empirical literature on differences in judicial behavior, which has focused on party, gender, and race. That literature has focused mainly on judges' votes and found partisan differences in some ideologically salient case categories, racial differences in a few case categories, and gender differences primarily in sex discrimination cases. In Part III we lay out our research questions. We focus on partisan differences as well as intraparty differences with respect to race and gender. In Part IV we present our data and research design, which are unique within the literature. Part V presents our primary results. In Part VI we discuss our findings and some of the

that Black, Hispanic, and female Republican district judges were more liberal (as measured by campaign contributions) than their White male Republican counterparts. *See* Maya Sen, *Diversity, Qualifications, and Ideology: How Female and Minority Judges Have Changed, or Not Changed, over Time*, 2017 WIS. L. REV. 367, 394 tbl.4 (2017) (finding preconfirmation ideological differences among district court judges within party based on race and gender).

16. For a similar finding in terms of how Supreme Court justices approach oral argument, see Tonja Jacobi & Matthew Sag, *The New Oral Argument: Justices as Advocates*, 94 NOTRE DAME L. REV. 1161, 1173–77 (2019) (discussing judicial polarization).

interesting questions they present, such as the greater partisan divergence for treatments of opinions written by Democratic appointees than for those written by Republican appointees—a difference that is consistent with an accelerating conservative shift among Republican appointees as opposed to a steady liberal shift among Democratic appointees. A brief conclusion follows.

I. THE IMPORTANCE OF TREATMENTS OF EARLIER OPINIONS

The vast majority of the scholarship on judicial behavior focuses on merits votes.[17] Federal appellate judges' votes to affirm or reverse a lower court's decision are obviously important. Which party prevailed is the most concrete outcome of an appeal, and the one that likely matters most to the parties in the case. Judges' votes both help to shape the law and reveal valuable information about the judges' preferences. But votes alone are a fairly crude metric.

At the outset, it bears noting that studies of votes rely on contestable (and contested) ideological coding of how conservative or liberal a given decision is.[18] For example, should *Gonzales v. Raich*[19] (upholding Congress's authority to criminalize marijuana production notwithstanding a state law allowing it) be coded as

17. *See infra* notes 44–46, 51–68 and accompanying text.

18. For arguments against the reliability of coding, see Hon. Harry T. Edwards & Michael A. Livermore, *Pitfalls of Empirical Studies that Attempt to Understand the Factors Affecting Appellate Decisionmaking*, 58 DUKE L.J. 1895, 1925 (2009):

> [I]t is very difficult to characterize many case outcomes. For example, the general rights embraced by freedom of religion and freedom of expression sometimes conflict with the exercise of other rights; it may not be clear how presumed liberal or conservative judges should be expected to vote in such cases. Cases may be disposed of on procedural grounds that are essentially nonideological, leading to coding errors when the outcome must be coded as liberal or conservative. A court's interpretation of a statute may defy ideological description (*e.g.*, rate allocations in a matter before the Federal Energy Regulatory Commission, where the parties before the court are competing companies) . . . . [M]any appeals involve multiple, complex issues, thus making it impossible to describe the appellate court's disposition as liberal or conservative.

See also Carolyn Shapiro, *Coding Complexity: Bringing Law to the Empirical Analysis of the Supreme Court*, 60 HASTINGS L.J. 477, 480–81 (2009) (criticizing ideological coding of cases); Anna Harvey & Michael J. Woodruff, *Confirmation Bias in the United States Supreme Court Judicial Database*, 29 J.L. ECON. & ORG. 414, 415–21 (2013) (finding, as the title suggests, confirmation bias in the ideological coding of cases).

For arguments in favor of the reliability of coding, see, for example, Tracey E. George, *Developing a Positive Theory of Decisionmaking on U.S. Courts of Appeals*, 58 OHIO STATE L.J. 1635, 1673 n.129 (1998) (arguing in favor of ideological coding).

19. Gonzales v. Raich, 545 U.S. 1, 74 (2005).

conservative, because it upheld drug laws, or liberal, because it upheld congressional authority under the Commerce Clause?[20]

Even if a coding scheme could deal with cases like *Raich* in a principled way, a judge's understanding of what counts as a conservative or liberal outcome is almost certainly highly timebound. What was regarded as a conservative decision in 1975 is likely different from what was regarded as conservative in 2015.[21] This will create difficulties for studies that attempt to analyze the conservatism or liberalism of merits votes over time.

But there is a deeper problem with focusing on judges' votes: the law is shaped by reasoning and by precedent—courts' treatments of earlier cases. The most important part of a given majority opinion's reasoning is its articulation of the test or factors that lead the court to decide as it does. Lawyers and judges interpreting a given opinion will look first to the court's articulation of its holding. But other aspects of opinions are significant, and revealing, even though they are not as important as the holding. One of the other significant aspects of an opinion is its treatment of earlier cases. The treatment of earlier cases helps to shape the law and concomitantly helps to reveal judges' preferences.

An opinion's treatment of earlier cases is important for the law's development in two related ways. First, the treatment of earlier cases is an important element of an opinion's reasoning. Opinions follow the precedents they deem controlling, and overrule, question, criticize, limit, or distinguish the opinions they deem not controlling or poorly reasoned. In a common law system, treatments of earlier cases are the building blocks for the substance of new opinions. Second, and relatedly, the substantive treatments of earlier opinions help shape legal doctrines. If later opinions repeatedly criticize or question a given opinion, a lawyer would be foolish to blithely rely on that case. Conversely, the more a given

20. *See* Ernest A. Young, *Just Blowing Smoke? Politics, Doctrine, and the Federalist Revival After* Gonzales v. Raich, 2005 SUP. CT. REV. 1, 11–14 (discussing the difficulties in coding *Raich*, and arguing that ideological coding of case outcomes is fraught with difficulties and that "the different factors used to code a case as 'conservative' or 'liberal' may cut in different directions within the confines of a single case").

21. To foreshadow one of our key results, such a change in perception as to what counts as conservative or liberal is consistent with our data and the associated explanation we put forth *infra* in Section VI.A.

case is followed, the stronger its precedential authority becomes.[22] If no court follows the reasoning of a given opinion, then that opinion does not shape the law. Thus, when a court follows a given opinion or diminishes its significance, that treatment not only constitutes part of the court's reasoning but also sends a signal to future courts about which opinions merit following and which do not.

This leads to an expectation that judges who are hostile to an earlier opinion on ideological (or other) grounds will be less likely to explicitly follow that opinion, and more likely to diminish its significance, and that the opposite will be true for judges who are supportive of an earlier opinion. Other aspects of opinion writing might be more important to judges than the treatment of earlier cases, but judges of course understand the significance of their treatments of earlier cases in shaping the law. So when it comes to responding to earlier cases, we would expect judges to be more likely to cast doubt on the opinions to which they are hostile and to treat as guiding precedents the opinions with which they agree.

Votes are not only significant but also more readily identifiable than are treatments of earlier cases. Insofar as judges want to move the law in their preferred ideological direction without attracting much notice (e.g., to avoid other judges on the panel or in the circuit objecting), we might expect the lower profile decision of how to treat a case to vary more with judge ideology than the higher profile decision of how to vote in a case. But the opposite seems at least as likely—that judges' greater focus on votes than on treatment of earlier cases will lead them to vote ideologically more than they treat earlier cases ideologically. The larger point is that, insofar as judicial behavior is correlated with ideology, we would expect that

22. See THOMAS G. HANSFORD & JAMES F. SPRIGGS II, THE POLITICS OF PRECEDENT ON THE U.S. SUPREME COURT 42–49 (2006), which develops a theory of how Supreme Court justices develop doctrine that measures opinions' positive and negative treatments and then empirically tests that theory. A key part of the Hansford and Spriggs theory is that a precedent has an amount of "vitality," or legal authority, that varies over time. Positive treatments of a precedent increase the vitality of that case while negative citations diminish vitality. According to Hansford and Spriggs, the ideological predispositions of the justices are moderated by the vitality of relevant precedents. A justice who prefers an outcome that is inconsistent with a vital key precedent will find it difficult to reach that outcome until the legal authority of the case has been chipped away over time via negative treatments. *See also* text accompanying note 48.

ideology would be associated with both voting and the treatment of earlier cases.

This raises the question of how to measure judicial treatment of earlier opinions. Citations are an obvious source. The most extensive study of judges' responses to their colleagues' opinions focuses on citations, on the theory that opinions cite to cases the author thinks are important.[23] But citations alone are a crude measure because they do not capture the nature of a court's treatment of an earlier opinion. Some citations occur in the context of a court relying on, and thus following, an earlier opinion. Some citations are negative (e.g., criticizing or questioning a precedent). And some may not be significant. A bare citation in a string of citations with no accompanying discussion of the cases does not provide much information.

To classify and measure the substantive treatment of earlier opinions in each majority opinion, we rely on *Shepard's*, an approach that has become standard in the literature.[24] *Shepard's* is a

23. *See* Stephen J. Choi & G. Mitu Gulati, *Bias in Judicial Citations: A Window into the Behavior of Judges*, 37 J. LEGAL STUD. 87, 94 (2008).

24. *See* Ryan C. Black & James F. Spriggs II, *An Empirical Analysis of the Length of U.S. Supreme Court Opinions*, 45 HOUS. L. REV. 621, 669–76 (2008); Chad Westerland, Jeffrey A. Segal, Lee Epstein, Charles M. Cameron & Scott Comparato, *Strategic Defiance and Compliance in the U.S. Courts of Appeals*, 54 AM. J. POL. SCI. 891, 896–98 (2010); James F. Spriggs II & Thomas G. Hansford, *Explaining the Overruling of U.S. Supreme Court Precedent*, 63 J. POL. 1091, 1097, 1100 (2001) [hereinafter Spriggs & Hansford, *Explaining the Overruling*]; Pamela C. Corley & Justin Wedeking, *The (Dis)Advantage of Certainty: The Importance of Certainty in Language*, 48 L. & SOC. REV. 35, 41–42, 47–48 (2014); HANSFORD & SPRIGGS, *supra* note 22, at 43–62; James F. Spriggs II & Thomas G. Hansford, *The U.S. Supreme Court's Incorporation and Interpretation of Precedent*, 36 L. & SOC. REV. 139, 146–47 (2002) [hereinafter Spriggs & Hansford, *U.S. Supreme Court's Incorporations*]; Christina L. Boyd & James F. Spriggs II, *An Examination of Strategic Anticipation of Appellate Court Preferences by Federal District Court Judges*, 29 WASH. U. J. L. & POL'Y 37, 63–67 (2009); Ryan C. Black & James F. Spriggs II, *The Citation and Depreciation of U.S. Supreme Court Precedent*, 10 J. EMPIRICAL LEGAL STUD. 325, 335 (2013); Frank B. Cross, *The Ideology of Supreme Court Opinions and Citations*, 97 IOWA L. REV. 693, 723–24 (2012); Stuart Minor Benjamin & Georg Vanberg, *Judicial Retirements and the Staying Power of U.S. Supreme Court Decisions*, 13 J. EMPIRICAL LEGAL STUD. 5, 10–11 (2016); Jeffrey Budziak, *The Effect of Visiting Judges on the Treatment of Legal Policy in the U.S. Courts of Appeals*, 38 JUST. SYS. J. 348, 349 (2017); Thomas G. Hansford, James F. Spriggs II & Anthony A. Stenger, *The Information Dynamics of Vertical Stare Decisis*, 75 J. POL. 894, 898 (2013); Lee Epstein, William M. Landes & Adam Liptak, *The Decision to Depart (or Not) from Constitutional Precedent: An Empirical Study of the Roberts Court*, 90 N.Y.U. L. REV. 1115, 1133 (2015); Ali S. Masood & Benjamin J. Kassow, *The Sum of its Parts: How Supreme Court Justices Disparately Shape Attention to Their Opinions*, 101 SOC. SCI. Q. 842, 853 (2020); Sara C. Benesh & Malia Reddick, *Overruled: An Event History Analysis of Lower Court Reaction to Supreme Court Alteration of Precedent*, 64 J. POL. 534, 540–41 (2002); Matthew P. Hitt, *Measuring Precedent in a Judicial Hierarchy*, 50 L. & SOC'Y REV. 57, 66–67 (2016); Michael P. Fix, Justin T. Kingsland &

widely used commercial legal research service that employs attorneys to examine every state and federal court opinion and code the content of every citation within each opinion. Sometimes a judge will cite an opinion in a way that seems to give no useful information about the citing judge's substantive treatment of that opinion (for instance, a string of citations with no discussion). If a citation refers to a case but has no meaningful substantive reaction to it, *Shepard's* does not put the citation into a substantive category. Such bare citations constitute the majority of citations.[25] *Shepard's* classifies citations that are accompanied by a substantive treatment of an opinion (i.e., a discussion of and substantive response to an opinion rather than a mere mention of it) into the following main categories: overruled, questioned, limited, criticized, distinguished, explained, harmonized, paralleled, and followed.[26] *Shepard's* characterizes overruled, questioned, limited, criticized, and

Matthew D. Montgomery, *The Complexities of State Court Compliance with U.S. Supreme Court Precedent*, 38 JUST. SYS. J. 149, 155–56 (2017); Scott D. McClurg & Scott A. Comparato, *Rebellious or Just Misunderstood?: Assessing Measures of Lower Court Compliance with U.S. Supreme Court Precedent*, RESEARCHGATE, 12–17 (Jan. 2004) (unpublished manuscript), (https://www.researchgate.net/profile/Scott-Mcclurg/publication/228504744_Rebellious_or_Just_Misunderstood_Assessing_Measures_of_Lower_Court_Compliance_with_US_Supreme_Court_Precedent/links/0912f50f41e34e92d5000000/Rebellious-or-Just-Misunderstood-Assessing-Measures-of-Lower-Court-Compliance-with-US-Supreme-Court-Precedent.pdf); Neal Devins & David Klein, *The Vanishing Common Law Judge?*, 165 U. PA. L. REV. 595, 605 (2017); Michael J. Nelson & Rachael K. Hinkle, *Crafting the Law: How Opinion Content Influences Legal Development*, 39 JUST. SYS. J. 97, 103–05 (2018); Joshua Boston, *Strategic Opinion Language on the U.S. Courts of Appeals*, 8 J.L. & CTS. 1, 8 (2020); Rachael K. Hinkle, *Strategic Anticipation of En Banc Review in the U.S. Courts of Appeals*, 50 L. & SOC'Y REV. 383, 393–94, n.3 (2016); Linda L. Berger & Eric C. Nystrom, *"Remarkable Influence": The Unexpected Importance of Justice Scalia's Deceptively Unanimous and Contested Majority Opinions*, 20 J. APP. PRAC. PROCESS 233, 252–54 (2019); Benjamin Kassow, The Impact of Ideology and Attorneys on Precedent Usage: An Analysis of State High Courts 99–101 (Jan. 1, 2013) (Ph.D. dissertation, University of South Carolina), (https://scholarcommons.sc.edu/etd/2571); Robert C. Wigton, *What Does It Take to Overrule? An Analysis of Supreme Court Overrulings and the Doctrine of Stare Decisis*, 18 LEGAL STUD. F. 3, 4 (1994); Michael C. Gizzi & R. Craig Curtis, *The Impact of* Arizona v. Gant *on Search and Seizure Law as Applied to Vehicle Searches*, 1 U. DENV. CRIM. L. REV. 30, 40–41 (2011); Rachael K. Hinkle, *Panel Effects and Opinion Crafting in the U.S. Courts of Appeals*, 5 J.L. & CTS. 313, 323 (2017) [hereinafter Hinkle, *Panel Effects*]; Pamela C. Corley, *Uncertain Precedent: Circuit Court Responses to Supreme Court Plurality Opinions*, 37 AM. POL. RSCH. 30, 36–37 (2009); Stuart Minor Benjamin & Bruce A. Desmarais, *Standing the Test of Time: The Breadth of Majority Coalitions and the Fate of U.S. Supreme Court Precedents*, 4 J. LEGAL ANALYSIS 445, 451–53 (2012).

25. *See* James F. Spriggs II & Thomas G. Hansford, *Measuring Legal Change: The Reliability and Validity of* Shepard's Citations, 53 POL. RSCH. Q. 327, 329 n.2 (2000).

26. *Shepard's* identifies these treatments as "overruled by," "questioned by," etc. In this Article we generally drop the "by" simply to avoid wordiness.

distinguished as negative treatments. More than 99.99% of the positive treatments in our dataset are followed, so in this Article we treat the following of an opinion and a positive treatment as synonymous and use the terms interchangeably.[27] Explained and harmonized are neutral treatments. They are rare and, as neutral treatments, are not indicative of a positive or negative response. Because we want to focus on substantive responses to other appellate opinions—that is, circuit court treatments that indicate some level of support or non-support—neutral responses are not of interest.[28]

The reliability of *Shepard's* treatments has been rigorously studied. Most notably, James Spriggs and Thomas Hansford undertook a careful study to measure the reliability of *Shepard's*. They took a stratified random sample of Supreme Court opinions citing earlier Supreme Court cases, yielding 602 citing opinions, and they coded all the citing opinions according to the coding rules in the *Shepard's* training manual. They found high levels of agreement between their coding and *Shepard's* coding.[29]

27. There are no instances of "paralleled" in our dataset. There are two other positive treatments, but they are quite rare, constituting less than .01% of the positive treatments: "extended by" (10 of the 648,226 treatments in our dataset) and "valid by" (1 of the treatments in our dataset).

Because "follow" is a verb, we generally use "follow" for the verb form and "positive treatment" for the noun form. Nothing substantive turns on this difference in terminology—we use it simply for clarity and ease of exposition.

28. *Shepard's* does not code the content of dissents, for good reason: by definition, substantive discussions in dissents do not represent the views of the majority and thus are not precedential. *Shepard's* does code concurrences, but it codes their treatments as neutral. *See, e.g.*, LEXISNEXIS, *SHEPARD'S* EDITORIAL PHRASES—ALPHABETICAL LIST 33, http://www.lexisnexis.com/pdf/lexis-advance/Shepards-Editorial-Phrases-Alphabetical-List.pdf (last visited Oct. 2, 2023) (listing "Criticized" in a majority opinion as a negative treatment, but "Criticized in Concurring Opinion" as a neutral treatment, noting for the concurrence that it "may not have the authority to materially affect its precedential value"). Again, by definition the substantive discussions in concurrences are not part of the majority opinion. That said, on some occasions a concurrence by a judge necessary to form the majority may contain an influential substantive discussion of an earlier case. *Shepard's* does not attempt to determine which concurrences may have some force and thus arguably might merit designation as positive or negative, for the apparent reason that such determinations are highly debatable. In this way, *Shepard's* may not code as positive or negative some treatments in concurrences that arguably are at least mildly positive or negative. We have no reason to believe that the absence of such information biases *Shepard's* results, and no other studies have so suggested or found. *See* Benjamin & Desmarais, *supra* note 24, at 7.

29. *See* Spriggs & Hansford, *supra* note 25, at 333–34.

There are strong reasons to believe that *Shepard's* treatments are valid measures. *Shepard's* definitions comport with judges' and lawyers' understanding of the substantive treatments. The specification of "followed" is illustrative. *Shepard's* defines "followed" as "[t]he citing opinion relies on the case you are Shepardizing as controlling or persuasive authority."[30] *Shepard's* created a training manual for the lawyers who code citations, with thirteen single-spaced pages devoted to laying out detailed coding rules for the treatment categories. According to the manual, "followed" (which the manual denotes with an "f") entails a case the citing opinion "relied on as controlling authority. The majority opinion in the [citing case] has expressly relied on the cited case as precedent on which to base its decision. The citing opinion must in some firm way refer to the cited case as compelling precedent."[31] The manual adds that "[a] mere 'going-along' with the cited case would not be sufficient for assigning a letter 'f.' Merely citing or quoting, with nothing more, is not a sufficient expression of reliance to permit an 'f' (or any other letter, for that matter)."[32] The manual identifies the following as language meriting a "followed" designation: "We affirm on the authority of . . . , or on the teaching of . . . , or for the reasons stated in . . . or under the rationale of . . . ; [or] such a conclusion is required by . . . or governed by . . . ."[33] This definition and discussion capture lawyers' and judges' understanding of what it means to follow a case.[34] This is not surprising, given the large amounts of money that lawyers have paid for access to *Shepard's*. It has long been widely used by practicing attorneys and judges, indicating that legal professionals view *Shepard's* as providing legally relevant information.[35]

It bears noting that positive treatment is much more common than negative treatment. Indeed, in our dataset there are 451,277 followed treatments and 141,768 negative treatments. The reason for this difference seems reasonably straightforward. For *Shepard's*

30. *See* LEXISNEXIS, HOW TO SHEPARDIZE: YOUR GUIDE TO LEGAL RESEARCH USING SHEPARD'S CITATIONS 10 (on file with authors).

31. *See* SHEPARD'S COMPANY, SHEPARD'S CITATIONS IN-HOUSE TRAINING MANUAL 13 (1993) (unpublished manual) (on file with authors).

32. *Id.*

33. *Id.*

34. *See* Benjamin & Vanberg, *supra* note 24, at 14.

35. *See, e.g.*, J. MYRON JACOBSTEIN & ROY M. MERSKY, FUNDAMENTALS OF LEGAL RESEARCH, 261–89 (1977).

to assign a negative treatment to a particular discussion in an opinion, that opinion must be explicit in its criticism, questioning, etc. of the earlier opinion. Such explicit negativity about an earlier opinion is fairly aggressive and might be perceived by other judges as uncollegial. No circuit judge wants her own opinion to be treated negatively by her colleagues in the future, and she might concomitantly be hesitant to treat her colleagues' earlier opinions negatively. Negative treatments are thus costly and relatively unusual. Following an earlier opinion has all the opposite attributes. As we noted above, it bolsters the opinion.[36] *Shepard's* characterizes following as the positive treatment for a reason—in our precedential system, following an opinion is the central form of praise. So the costs of following an opinion are quite low. The only disadvantage of following an earlier opinion for a later panel would arise if the later panel did not in fact want to provide support for the earlier opinion. If the later panel found the earlier opinion objectionable, then, and only then, would it have an incentive to avoid following that earlier opinion. Indeed, this last point highlights why examining positive treatments can be so revealing: we would expect judges to be more inclined to follow opinions with which they agree and less inclined to follow opinions with which they do not.[37]

Are these *Shepard's* treatments reflective of judges' choices? There are two possible ways in which an opinion's discussion of a previous opinion might not reflect a judge's meaningful decision. One is that the judge may effectively have no choice in the matter. Most obviously, if there is only one precedent that directly controls the question at issue, then we would expect (or at least hope) that any judge would follow that precedent. Insofar as existing precedents constrain judges, one element of that constraint is that there are some cases that can have only one possible result, because that is what precedent demands. After *Roe v. Wade*[38] (and before

36. *See supra* note 22 and accompanying text.

37. In this Article we often use the word "citing" to refer to *Shepard's* treatments for the sake of streamlining some sentences. But, to be clear, what we are studying is *Shepard's* substantive treatments, and, as we have just discussed, those treatments are much more than a mere citation. So any references to citations in our data are referring to substantive treatments, and we use the terms "citing" and "treating" interchangeably.

38. Roe v. Wade, 410 U.S. 113 (1973).

Dobbs v. Jackson Women's Health Organization[39]), for example, any lower court would be compelled by *Roe* to invalidate a statute that criminalized all abortions.

But many cases are not so clearly constrained by binding legal precedent. To return to the abortion example, a flat ban on abortion was foreclosed by *Roe*, but restrictions on some abortions were subject to differing interpretations (as has been the case when courts of appeals have reviewed post-*Roe* abortion restrictions).

A second limit to the force of the argument may be more significant: there will rarely be only one precedent on which a court can rely. Consider a question like the standard applied for issuing a preliminary injunction or summary judgment. There are thousands of cases laying out a standard, and they often differ, even if only slightly, in the wording they use. Some wording is slightly more favorable to those seeking the injunction or summary judgment, and some is slightly less favorable. Judges can choose among them when deciding which case to follow in the articulation and application of standards for an injunction or summary judgment.

And even in situations where there is only one precedent directly on point for a given case, as soon as that case is decided there will be two cases that are directly on point. So the next panel confronting the same issue will have a choice among two relevant precedents, and the panel after that will have three, and so on. And given that each of those opinions will differ slightly from the others and will have different authors, judges will be able to make some choices in determining which opinions they follow.

The second possible way in which opinions' discussion of cases might not reflect meaningful decisions by judges is that judges may leave those decisions to their clerks. Insofar as judges defer to their clerks (or anyone else) in their opinions' discussions of earlier cases, the judges are not making the meaningful decisions and we should not expect to see the differences in treatment behavior that we hypothesize.

The intuition behind this second possibility is that judges often rely on their clerks for the first draft of a majority opinion and in particular may rely on their clerks for matters as mundane as which cases to cite. We think that this intuition likely has particular force with respect to string cites in which an opinion states a basic legal

39. Dobbs v. Jackson Women's Health Org., 142 S. Ct. 2228 (2022).

rule and then lists several cases that support that legal rule. It seems quite unlikely that judges choose (or even focus on) every case that they list in a string cite for a straightforward proposition. But, as we noted above, *Shepard's* does not code such citations as positive treatments. For *Shepard's* to code something as a treatment, the opinion must have a meaningful and significant discussion of the case—a bare citation does not count. As we also noted, in our precedent-based system discussions of earlier opinions, and decisions of which opinions to follow (and which to cast doubt on), are important elements of a given opinion. Judges are much more likely to make choices about crafting those elements than they are about what cases to list in a string cite.

But we recognize that judges may well defer to their clerks' choice of which cases to rely on (and thus follow for *Shepard's* purposes) and which cases to diminish (and thus overrule, criticize, question, distinguish, or limit for *Shepard's* purposes). Indeed, some judges in some opinions may well defer to their clerks on all aspects of an opinion.

This raises an important possible dampening effect. The possibilities of a single directly relevant precedent and the effect of clerks will tend to diminish the differences we are studying in this Article: the less that opinions reflect judges' choices, the less likely that there will be significant differences in the measures designed to capture those choices. Insofar as we find the differences we hypothesize, we find those differences despite the dampening impact of these possibilities.

We have no reason to believe that either of these possibilities would skew our data. As to clerks, some judges may focus heavily on ideology in choosing their clerks and choose clerks who are ideologically aligned with them, but that of course would be an accurate reflection of the judge's ideology.[40] For other judges, ideology may play no role in their choice of clerks. For those judges, the impact of clerks would be random and thus would mute

40. *See, e.g.*, Adam Bonica, Adam S. Chilton, Jacob Goldin, Kyle Rozema & Maya Sen, *Measuring Judicial Ideology Using Law Clerk Hiring*, 19 AM. L. & ECON. REV. 129, 146 (2017) ("[C]lerk ideologies provide a window into the ideology of the hiring judge . . . ."); Adam Bonica, Adam S. Chilton, Jacob Goldin, Kyle Rozema & Maya Sen, *The Political Ideologies of Law Clerks*, 19 AM. L. & ECON. REV. 96, 123–24, (2017) (finding evidence that judges hire clerks with similar ideologies).

the impact of judicial partisanship.[41] But what if, say, Republican appointees are more likely than Democratic appointees to choose clerks who are ideologically aligned with them? Under those circumstances, clerks' influence would push Democratic appointees toward the middle of the ideological spectrum while their Republican counterparts would be unchanged. That might affect attempts at measuring the absolute conservatism or liberalism of appellate opinions. But our focus is on the relative distance between Republican and Democratic appointees. We are not measuring whether Democratic or Republican appointees are more liberal or conservative than expected. We are simply measuring the divergence between Democratic and Republican appointees. If Democrats were less likely than Republicans to hire ideologically compatible clerks (or vice versa), this would tend to dampen partisan differences (the fewer the clerks hired for ideological compatibility, the less a clerk effect will lead to ideological differences in how judges treat earlier opinions). Relatedly, if clerks were more likely to be politically moderate than the judges for whom they clerk, that might mute the differences in behavior between Democratic and Republican appointees and thus mute any differences based on the partisanship of the judges.

By contrast, if it were the case that Republican appointees systematically choose clerks who are more conservative than they, and that Democratic appointees systematically choose clerks more liberal than they, then this heterogeneity among clerks could increase the partisan differences we see between Democratic and Republican appointees. But in such circumstances that heterogeneity would be a function of judges' decisions to choose clerks who are more ideologically extreme than they are. That is, the heterogeneity would reflect the judges' ideological disposition to have clerks who are more ideological than the judges themselves are. Differences in clerks would be attributable to the judges who hired them. Similar points apply to the race and gender of clerks. Insofar as some judges are disproportionately likely to hire clerks whose race or gender matches their own, the role of the clerk's race or gender plays a similar role to that of the judge's race or gender. Conversely, insofar

41. *See* Jeremy D. Fogel, Mary S. Hoopes & Goodwin Liu, *Law Clerk Selection and Diversity: Insights from Fifty Sitting Judges of the Federal Courts of Appeals*, 137 HARV. L. REV. (forthcoming Nov. 2023) (finding that "[m]ost judges disclaim any interest in ideological alignment when hiring clerks[,]" though of course this is self-reported).

as other judges hire without regard to race or gender, any impact the clerks have might reduce the measured effect of judges' race and gender on how they treat earlier cases.

As to precedents directly on point, there is also no reason to believe there is any skew that would affect our analysis. We have no reason to believe that there is any difference in the likelihood of Democratic versus Republican, White versus Hispanic versus Black, or female versus male judges to write the sorts of opinions that are likely to be followed. But even if that were true, then presumably Democratic and Republican, White, Hispanic, and Black, and female and male judges would follow those opinions to a similar degree. If, say, Republican appointees are more likely to write opinions that merit being relied on, then presumably that reliance would be across the board and there would be no differences in the likelihood of particular categories of later judges to rely on them. And if the response to that last point is that later Republican appointees (to stick with the example) are more likely than Democratic appointees to find merit in earlier opinions by their Republican colleagues—well, that is exactly what we are trying to measure. That would not be a skewing of our data; it would be a confirmation of our hypothesis.

II. THE EMPIRICAL LITERATURE ON JUDICIAL BEHAVIOR—PARTY, GENDER, AND RACE

One of the central questions at the intersection of law and political science is to what extent judges' personal characteristics influence their judicial behavior. Can we learn anything from examining the relationship between some personal attributes of judges and what they do on the bench?

Many studies examine how the individual attributes of judges correlate with judges' behavior. Many of these studies focus on differences between Republican and Democratic appointees' judicial behavior, but some address the relationship between race or gender and differences in judicial behavior. These studies generally look directly at individual voting differences and are thus easily interpretable. Our analyses are similar, but we look at differences in majority opinion authors' treatments of earlier opinions.

A. Political Party

The most commonly studied characteristic of judges is the political party of the President who most recently nominated a given judge.[42] The underlying theory is that judicial behavior may be influenced by ideology as revealed by the President's party. The key elements of that theory are fairly straightforward:

- Democratic and Republican Presidents diverge ideologically;
- Party and ideology are not perfectly correlated, but there is a strong relationship between the two;
- Presidents choose circuit court nominees with whom they are ideologically compatible;
- Legal doctrine may impose meaningful constraints but often leaves room for judicial decisions that are not determined by legal doctrine and can be influenced by ideology; and
- One of the things judges seek to achieve (indeed, one of the reasons to want to be a judge) is to help move the law in a positive direction, and a given judge's definition of "positive" will be correlated with the judge's ideology, with the result that judges will thus want to push the law in an ideological direction (even though they may conceptualize the direction as "positive" rather than ideological).[43]

In light of the importance of determining the impact of the President's party on judicial behavior, many empirical studies have

42. *See, e.g.*, Pauline T. Kim, *Deliberation and Strategy on the United States Courts of Appeals: An Empirical Exploration of Panel Effects*, 157 U. PA. L. REV. 1319, 1327 (2009) (noting the party of the most recent President to nominate a judge is the standard practice for identifying the ideology of a judge); SUNSTEIN ET AL., *supra* note 7, at 5–7 (using the President's party as the measure of judicial ideology); Adam B. Cox & Thomas J. Miles, *Judging the Voting Rights Act*, 108 COLUM L. REV. 1, 3 (2008) (same); Richard L. Revesz, *Environmental Regulation, Ideology, and the D.C. Circuit*, 83 VA. L. REV. 1717, 1718–19 (1997) (same); *see also* Daniel R. Pinello, *Linking Party to Judicial Ideology in American Courts: A Meta-analysis*, 20 JUST. SYS. J. 219, 221 (1999) (providing "a compendium of empirical undertakings connecting party ID with judicial ideology"). Note that because every President has been a Republican or Democrat, dividing judges into Republican and Democratic categories captures all judges.

43. *See, e.g.*, SUNSTEIN ET AL., *supra* note 7, at 3, 13, 22–27. Some scholars associated with the attitudinalist model argue that legal doctrine poses little or no constraint, and that judges decide cases primarily (and sometimes exclusively) based on their ideology, but one need not subscribe to that view to posit that ideology likely plays *some* role. *See, e.g.*, JEFFREY A. SEGAL & HAROLD J. SPAETH, THE SUPREME COURT AND THE ATTITUDINAL MODEL REVISITED 37–42 (2002).

attempted to measure that impact. Most such studies—indeed, most studies of judicial behavior—focus on judges' votes as the relevant behavior to be measured. The most extensive study of judicial voting in the U.S. Courts of Appeals was conducted by Cass Sunstein, David Schkade, Lisa Ellman, and Andres Sawicki. They focused on judges' votes in published opinions in twenty-four issue areas that might be expected to have a partisan valence, such as environmental law, sex discrimination, sexual harassment, disability discrimination, and campaign finance, and found partisan differences in fifteen of them.[44] Other, less comprehensive studies have found differences in judges' votes in ideologically salient areas like voting rights, affirmative action, and employment discrimination.[45] But other studies have found an absence of differences in some politically charged areas (such as abortion and capital punishment).[46]

Votes are not the only outcome that can be studied. A natural alternative is to examine citation practices. These studies have generally focused on the Supreme Court and have found differences in citation behavior (relying on *Shepard's*) that comport with differences in judicial voting. For instance, Hansford and Spriggs find that Supreme Court justices are more likely to

44. *See* SUNSTEIN ET AL., *supra* note 7, at 26–27.

45. *See* Cox & Miles, *supra* note 42, at 48; Jonathan P. Kastellec, *Racial Diversity and Judicial Influence on Appellate Courts*, 57 AM. J. POL. SCI. 167, 173 (2013) (finding that Democratic appointees were more likely than Republican appointees to vote in favor of affirmative action); Sean Farhang & Gregory Wawro, *Institutional Dynamics on the U.S. Court of Appeals: Minority Representation Under Panel Decision Making*, 20 J.L. ECON. & ORG. 299, 314 n.10, 320 (2004) (finding that judges appointed by more conservative Presidents were less likely to find for the plaintiff in employment discrimination cases); *see also* Richard L. Revesz, *Environmental Regulation, Ideology, and the D.C. Circuit*, 83 VA. L. REV. 1717, 1742–43 (1997) (finding that the party of the appointing President is associated with differences in judges' votes in environmental cases on the D.C. Circuit).

46. *See* Thomas J. Miles & Cass R. Sunstein, *The New Legal Realism*, 75 U. CHI. L. REV. 831, 839 (2008). There have been varying findings of partisan differences in religious freedom cases. An early study found partisan differences on voting in Free Exercise claims, particularly where the subjects at issue were politically charged, such as religious accommodations for children in school, with judges appointed by Republican Presidents being more likely to vote in favor of such accommodations. *See* Gregory C. Sisk, Michael Heise & Andrew P. Morriss, *Searching for the Soul of Judicial Decisionmaking: An Empirical Study of Religious Freedom Decisions*, 65 OHIO STATE L. J. 491, 602 (2004). In contrast, a later study found no statistically significant partisan difference in Free Exercise claims. Both studies, however, found that Republicans were less likely to vote in favor of claimants in Establishment Clause cases. *See* Sepehr Shahshahani & Lawrence J. Liu, *Religion and Judging on the Federal Courts of Appeals*, 14 J. EMPIRICAL LEGAL STUD. 716, 731–32 (2017).

positively treat ideologically proximate precedents and negatively treat ideologically distant precedents.[47] They go on to show that these citation practices affect the "vitality" of the cited case, with positive treatments increasing vitality and negative treatments diminishing vitality.[48] Chad Westerland et al. find that a lower court's propensity to positively treat a Supreme Court opinion depends on the ideological distance between the enacting Supreme Court and the current Supreme Court, but not on the ideological distance between the lower court and either the enacting or current Supreme Court.[49] And Frank Cross et al. find that the ideological heterogeneity of the Supreme Court majority coalition is significantly, albeit moderately, associated with citation practices. Specifically, they find that more ideologically heterogeneous Supreme Court majority coalitions cite more opinions than ideologically homogeneous coalitions and the opinions that are cited have greater network centrality.[50]

B. Gender and Race

Partisanship is a significant attribute that might affect judicial behavior, but it is not the only one. After all, parties, and cohorts of judges within those parties, have variation within them. After partisanship, the two most prominent attributes assessed in studies of judicial behavior are judges' gender and race. These studies raise the possibility that within parties gender and race may shed light on judicial behavior.

1. Gender

Previous studies of racial or gender differences in judicial behavior have generally focused on judges' votes in areas of law thought to activate gender and racial identities.[51] In the gender

47. *See* HANSFORD & SPRIGGS, *supra* note 22, at 94.

48. *Id.* For a discussion of vitality, see *supra* note 22.

49. *See* Westerland et al., *supra* note 24, at 905.

50. Frank B. Cross, James F. Spriggs II, Timothy R. Johnson & Paul J. Wahlbeck, *Citations in the U.S. Supreme Court: An Empirical Study of Their Use and Significance*, 2010 U. ILL. L. REV. 489, 551.

51. The most notable exceptions are two articles by Stephen Burbank and Sean Farhang finding some race and gender effects for class certification decisions and some gender effects for motions to dismiss for failure to state a claim after *Twombly* and *Iqbal*. *See* Stephen B. Burbank & Sean Farhang, *Politics, Identity, and Class Certification on the U.S. Courts*

context, Christina Boyd, Lee Epstein, and Andrew Martin identified these areas of law as case types "on which female judges may possess valuable expertise, experience, or information" or that are issues of concern to women broadly speaking, such as sex harassment and sexual discrimination.[52] Such cases are often called gender-coded or gender-salient, and we use the latter term.[53] The idea is that in specific kinds of cases, judges may have a particular understanding of and sensitivity to particular issues, perhaps flowing from their expertise and lived experiences.[54]

Studies have found relatively few gender differences in voting. Boyd, Epstein, and Martin identified thirteen gender-salient categories of cases and found gender differences in only one: sex discrimination in employment.[55] Similarly, Susan Haire and Laura Moyer found that "[m]en and women on the bench are quite similar in their voting behavior, with one exception: cases involving sex discrimination."[56] Sarah Westergren, meanwhile, found that any

of Appeals, 119 MICH. L. REV. 231, 231 (2020) [hereinafter Burbank & Farhang, *Class Certification*] (finding that "the presence of one African American on a panel, and the presence of two women (but not one), is associated with procertification outcomes."); Stephen B. Burbank & Sean Farhang, *Politics, Identity, and Pleading Decisions on the U.S. Courts of Appeals* 169 U. PA. L. REV. 2225, 2226 (2021) [hereinafter Burbank & Farhang, *Pleading Decisions*] (finding that in precedential cases "panels with one woman were more likely to decide precedential other civil rights claims in favor of plaintiffs, and that panels with two women (but not one) were more likely to do so in non-civil rights claims.").

52. *See* Christina L. Boyd, Lee Epstein & Andrew D. Martin, *Untangling the Causal Effects of Sex on Judging*, 54 AM. J. POL. SCI. 389, 391 (2010).

53. *See id.* Gender-salient (and race-salient) categories of cases are a subset of the ideologically salient categories we identify. *See infra* note 82 & Appendix A1.

54. There are various theories behind this proposition: that female judges bring a unique knowledge base in key areas like sex discrimination based on their experiences; that men and women think, communicate, and view the world differently from one another; and/or that judges serve as representatives of their group and work to advance their group's interests. *See, e.g.*, Christina Boyd, *Representation on the Courts? The Effects of Trial Judges' Sex and Race*, 69 POL. RES. Q. 788, 789–90 (2016); *infra* notes 81–82 and accompanying text.

55. *See* Boyd, Epstein & Martin, *supra* note 52, at 389 (noting that among the 13 areas of law with gender salience, they observed gender differences only for sex discrimination in employment claims).

56. SUSAN B. HAIRE & LAURA P. MOYER, DIVERSITY MATTERS: JUDICIAL POLICY MAKING IN THE U.S. COURTS OF APPEALS 53–54 (2015); *see also id.* at 48 (female judges "tend to decide cases similarly to their male colleagues" with a single exception: "Women judges are more likely to support plaintiffs in sex discrimination cases when compared to the votes of their male colleagues."). Interestingly, Haire and Moyer found that this difference in votes in sex discrimination cases is a function of age and experience: older cohorts of women and men voted differently in sex discrimination cases, but more recent cohorts did not. *See* Laura P. Moyer & Susan B. Haire, *Trailblazers and Those that Followed: Personal Experiences, Gender, and*

gender differences in sex discrimination disappeared once she controlled for judges' partisanship.[57] And, strikingly, Jennifer Segal, examining a range of gender-salient cases decided by Clinton-appointed district court judges, found that "there are gender differences in cases involving women's issues, yet it is male judges who are more supportive of these claims."[58]

A few studies have looked beyond gender-salient case types to examine gender differences in areas that are ideological but have no obvious gender salience (such as voting rights and religious liberty), but those studies have not found gender differences in judges' votes in such cases.[59] For example, Sunstein, Schkade, Ellman, and Sawicki found no gender differences in a broad range of ideologically salient cases, and Haire and Moyer aggregated all case types and found no statistically significant gender differences.[60]

Judicial Empathy, 49 L. & SOC'Y REV. 665, 668 (2015). For a discussion of the hypotheses that flow from a possible trailblazer effect, see *infra* notes 93–94 and accompanying text. It also bears noting that in an earlier, much smaller study, Jennifer Peresie also found panel and judge effects in votes in sexual harassment cases. Jennifer L. Peresie, Note, *Female Judges Matter: Gender and Collegial Decisionmaking in the Federal Appellate Courts*, 114 YALE L.J. 1759, 1776 (2005) ("[I]n Title VII sexual harassment and sex discrimination cases, . . . a judge's gender and the gender composition of the panel mattered to a judge's decision.").

57. *See* Sarah Westergren, *Gender Effects in the Courts of Appeals Revisited: The Data Since 1994*, 92 GEO. L.J. 689, 703 (2004) (finding no statistically significant effect of gender on judges' votes in sex discrimination cases, and that "any gender effect appears to be intertwined with the effect of political party affiliation of the appointing president"); *see also* Carol T. Kulik, Elissa L. Perry & Molly B. Pepper, *Here Comes the Judge: The Influence of Judge Personal Characteristics on Federal Sexual Harassment Case Outcomes*, 27 L. & HUM. BEHAV. 69, 80–81 (2003) (finding no gender effect on sexual harassment cases).

58. Jennifer A. Segal, *Representative Decision Making on the Federal Bench: Clinton's District Court Appointees*, 53 POL. RES. Q. 137, 144 (2000).

59. *See, e.g.*, Cox & Miles, *supra* note 42, at 43 (finding no statistically significant effect of gender on judges' voting patterns in voting rights cases); Sisk et al., *supra* note 46, at 593 (finding no statistically significant effect of gender on judges' voting patterns in religious liberty cases); Kastellec, *supra* note 45, at 178 (finding no statistically significant effect of gender on judges voting in affirmative action cases regarding race); Kenneth L. Manning, Bruce A. Carroll & Robert A. Carp, *Does Age Matter? Judicial Decision Making in Age Discrimination Cases*, 85 SOC. SCI. Q. 1, 12 tbl.2 (2004) (finding no statistically significant gender differences in judges' voting in age discrimination cases); Orley Ashenfelter, Theodore Eisenberg & Stewart J. Schwab, *Politics and the Judiciary: The Influence of Judicial Background on Case Outcomes*, 24 J. LEGAL STUD. 257, 262 (1995) (finding only "modest" effects of gender on judges' voting in civil rights cases).

60. SUNSTEIN ET AL., *supra* note 7, at 167, 171, 185, 197; HAIRE & MOYER, *supra* note 56, at 47–48 (finding no evidence of female judges voting differently from their colleagues when aggregating all case types).

Thus, the studies have found very limited voting differences based on gender. As Christina Boyd and Adam Rutkowski put it, "[a] relatively large number of empirical studies . . . have failed to find evidence that female and male judges decide cases differently from one another, particularly outside of issue areas that are not closely related to 'women's issues' like sex discrimination."[61]

2. Race

Studies have found racial voting differences in some race-salient case types but not in broader categories of cases. For instance, Jonathan Kastellec found racial voting differences in affirmative action and death penalty cases. Looking at all affirmative action decisions regarding race in the U.S. Courts of Appeals between 1971 and 2008, Kastellec found that Black judges were more likely to support affirmative action programs.[62] In death penalty cases, Kastellec found that adding a Black judge to a non-Black panel significantly increased the chances of granting relief to defendants on death row when the defendant was Black.[63] Other studies have found similar differences in other race-salient cases, such as racial harassment, voting rights, and police misconduct cases.[64] The underlying theory with respect to race-salient cases is that non-White judges will approach these issues differently than their White counterparts because of their life experiences and views given the long history of racial discrimination in the United States.[65]

61. Christina L. Boyd & Adam G. Rutkowski, *Judicial Behavior in Disability Cases: Do Judge Sex and Race Matter?*, 8 POL., GRPS., & IDENTITIES 834, 837–38 (2020). One theory behind this outcome suggests that all judges, regardless of background, are so influenced by their training prior to taking the bench, and constrained by judicial norms and practices, that any differences from their background are offset and have no systematic impacts on their behavior. *See* Boyd, *supra* note 54, at 790.

62. Kastellec, *supra* note 45, at 179; *see also*, Peresie, *supra* note 56, at 1774, 1776 (same).

63. Jonathan P. Kastellec, *Race, Context, and Judging on the Courts of Appeals: Race-Based Panel Effects in Death Penalty Cases*, JUST. SYS. J., Nov. 11, 2020, at 410.

64. *See, e.g.*, Pat K. Chew & Robert E. Kelley, *Myth of the Color-Blind Judge: An Empirical Analysis of Racial Harassment Cases*, 86 WASH. U. L. REV. 1117, 1134 (2009) (finding that plaintiffs in racial harassment cases are more than twice as likely to succeed under a Black judge than a White judge); Cox & Miles, *supra* note 42, at 30, 43 (finding that a judge's race is associated with higher likelihood of voting in favor of liability in voting rights cases on the U.S. Courts of Appeals); Nancy Scherer, *Blacks on the Bench*, 119 POL. SCI. Q. 655, 668 (2004) (finding a "statistically significant difference in the voting behavior between black and white judges" in police misconduct cases).

65. HAIRE & MOYER, *supra* note 56, at 25, 32–33.

As with gender, studies have generally not found racial differences when studying more broadly ideologically salient cases or when aggregating all case types together.[66] Thus for both gender and race, voting differences have been confined to a small subset of cases that are gender- or race-salient. Scholars have typically not found broader differences that comport with political ideology more generally.

One common, and significant, limitation of the studies looking for party, gender, and racial differences was that very few were able to control for year and circuit.[67] A major reason for this is that previous researchers have worked with limited datasets, usually containing hundreds of cases,[68] which makes statistical adjustment difficult.

66. *See id.*, at 32; SUNSTEIN ET AL., *supra* note 7, at 167, 171, 185, 197; Farhang & Wawro, *supra* note 45, at 321 ("[R]acial minority judges on the federal Court of Appeals in the period sampled do not hold views different from White judges on employment discrimination claims, as measured by case outcome."); *see also* HAIRE & MOYER, *supra* note 56, at 32 (finding that Black judges do not vote more liberally than White judges on the Courts of Appeals when data is pooled over many policy areas). *But see* Sisk et al., *supra* note 46, at 595–96 (finding that a judge's minority race is associated with a higher likelihood of voting in favor of plaintiffs alleging religious discrimination, and that minority judges are more willing to take non-mainstream approaches in religious freedom cases).

67. *See, e.g.*, Kastellec, *supra* note 45, at 172–73 (neither controlling by year nor circuit); Sisk et al., *supra* note 46, at 553–55 (same); HANSFORD & SPRIGGS, *supra* note 22, at 42–46 (same); Hinkle, *Panel Effects*, *supra* note 24, at 322, 324 (same); Boyd, *supra* note 54, at 792–93 (same); Cox & Miles, *supra* note 42, at 21–22, 25–26 (same); Manning et al., *supra* note 59, at 7–8 (same); Chew & Kelley, *supra* note 64, at 1138 (same). Some were able to control by either year or circuit. *See, e.g.*, Farhang & Wawro, *supra* note 45, at 315 (controlling for circuit); Peresie, *supra* note 56, at 1775–76 (same); Scherer, *supra* note 64, at 666 (controlling by region and defining some regions by circuit); Kulik et al., *supra* note 57, at 76, 80 (controlling by year); HAIRE & MOYER, *supra* note 56, at 159 (same). Very few were able to control for both. One of the few exceptions is Shahshahani & Liu, *supra* note 46, at 726 (controlling for both year and circuit).

68. *See, e.g.*, Kastellec, *supra* note 45, at 173 (studying a total of 182 cases between 1971 and 2008); Sisk et al., *supra* note 46, at 553 (studying 729 or fewer decisions depending on the model used); Boyd, *supra* note 54, at 793 (studying between 186 and 450 observations of judge voting); Cox & Miles, *supra* note 42, at 8 (studying 342 decisions); Manning et al., *supra* note 59, at 5 (studying 544 cases); Chew & Kelley, *supra* note 64, at 1138 (studying 428 cases); Farhang & Wawro, *supra* note 45, at 310 (studying 400 cases); Kulik et al., *supra* note 57, at 75 (studying 143 cases); Peresie, *supra* note 56, at 1767 (studying 556 cases); Scherer, *supra* note 64, at 672 (studying 550 cases). *But see* HANSFORD & SPRIGGS, *supra* note 22, at 51 (studying 6,363 cases); Hinkle, *Panel Effects*, *supra* note 24, at 322 (including 6,693 cases in its study); Shahshahani & Liu, *supra* note 46, at 721, 735–36 (studying 1,058 religious freedom cases and 2,100 cases not involving religion).

III. OUR RESEARCH QUESTIONS

The questions we want to examine in this Article are all elaborations of a simple inquiry: How do judge-specific characteristics relate to the substantive treatment of earlier opinions? To answer these questions, we focus on majority opinion authors and estimate differences in substantive treatment practices across different types of authors.[69] Importantly, in all our analyses we adjust for circuit, year, and circuit-year effects to get as close as possible to apples-to-apples comparisons that are descriptively informative about substantive treatments.[70] And, in light of the possibility of temporal changes in behavior, particularly in light of increases in measures of polarization among decisionmakers in the many years our data cover, we also evaluate changes over time. To do this, we split our data into four equally sized time periods (1974–84, 1985–95, 1996–2006, and 2007–17) and estimate partisan, racial, and gender differences within each time period.

A. Partisan Differences

The first opinion-author-specific attribute that we examine is partisanship as proxied by the appointing President's party.[71] More specifically, we ask: Are opinion authors of a given party more likely than authors from the opposite party to follow opinions written by fellow members of their party?

If judges are pervasively partisan, then we might expect to find statistically significant results if we look at all panels in all types of cases. After all, insofar as judges are deeply partisan, we might expect their partisanship to arise across the board. But such pervasive partisanship may seem somewhat unrealistic. Even those in the political branches find room to agree on some relatively less ideological matters.

69. We focus on opinion authors in light of the centrality of their role on the panel in crafting the discussion contained in the majority opinion. *See, e.g.*, Chad M. Oldfather, *Remedying Judicial Inactivism: Opinions as Informational Regulation*, 58 FLA. L. REV. 743, 794–95 (2006). In a separate paper we measure the extent to which there are partisan panel effects (i.e., whether the party of other panel members influences substantive treatments contained in the majority opinion).

70. *See* Fan Li, Alan M. Zaslavsky & Mary Beth Landrum, *Propensity Score Weighting with Multilevel Data*, 32 STAT. MED. 3373–87 (2013) (describing such average controlled differences).

71. *See supra* notes 42–43 and accompanying text.

It is not clear that there are any categories of cases that are totally nonideological. Even run-of-the-mill torts and contracts cases can have some political valence (e.g., perhaps the average Republican appointee is relatively more sympathetic to defendants in torts cases and to enforcing contracts than is the average Democratic appointee). But we might expect cases involving politically charged issues like campaign finance and affirmative action to induce more ideological behavior than ordinary torts and contracts cases. And in fact most of the empirical studies on partisan judicial behavior focus on case topics that are expected to be ideological and thus polarized on partisan grounds.[72]

Thus a narrower form of the hypothesis above would expect larger partisan differences within case topics that are most likely to be ideologically charged along partisan lines. We canvassed previous studies for the case topics they identified as more likely to divide judges along political ideology lines, a category of cases we refer to as ideologically salient.[73] We then identified all the Lexis topics that involved one of these case topics. That yielded thirty-eight Lexis case topics.[74] Note that some of these categories had relatively few cases (e.g., Establishment Clause and abrogation of state sovereign immunity).[75]

A focus on more ideological cases implicates the distinction between published and unpublished cases. For much of the period our data cover, circuit rules prohibited or at a minimum disfavored citation of unpublished cases.[76] Circuit rules allow the ruling panel

72. *See supra* notes 44–46 and accompanying text.

73. *See supra* notes 44–46 and accompanying text.

74. We list the thirty-eight topics in Appendix A1. We identified a thirty-ninth ideologically salient topic (federalism), but Lexis did not use federalism as a case topic header in any of the cases in our dataset.

75. Previous studies may have chosen these categories in part because the number of cases was small enough to allow them to address all the cases. Our dataset contains the entire universe of cases and we wanted to separate by circuit and year to isolate effects, so the small numbers in some categories made it extremely unlikely that we would find statistical significance. The error bars in the accompanying figures reflect this.

76. *See, e.g.*, In re Citation of Unpublished Opinions/Orders and Judgments, 151 F.R.D. 470 (Nov. 29, 1993) (replacing its prohibition on citation of unpublished opinions with the following rule: "Unpublished opinions and orders and judgments of this court are not binding precedents, except under the doctrines of law of the case, res judicata, and collateral estoppel. Citation of these unpublished decisions is not favored."); FED. R. APP. P. 32.1. ("A court may not prohibit or restrict the citation of federal judicial opinions, orders, judgments, or other written dispositions that have been: (i) designated as 'unpublished,' 'not for

to choose whether to publish an opinion based on its importance and precedential value.[77] The point of the unpublished designation is to allow judges to issue relatively insignificant opinions. Indeed, each year thousands of very short opinions (often only a paragraph or two long, sometimes only a single sentence) are issued, virtually all of which are unpublished. Some have suggested that judges have on occasion refrained from publishing a given opinion to diminish its significance.[78] This highlights that there is some discretion involved in the decision to publish an opinion. But there is no evidence of partisan, racial, or gender differences in decisions not to publish an opinion or to follow an unpublished opinion, so we have no reason to believe that the decision to publish or not publish a given opinion would affect our findings.[79]

---

publication,' 'non-precedential,' 'not precedent,' or the like; and (ii) issued on or after January 1, 2007.")

77. For example, as the Ninth Circuit outlines in its "CRITERIA FOR PUBLICATION:"

> A written, reasoned disposition shall be designated as an OPINION if it: (a) Establishes, alters, modifies or clarifies a rule of federal law, or (b) Calls attention to a rule of law that appears to have been generally overlooked, or (c) Criticizes existing law, or (d) Involves a legal or factual issue of unique interest or substantial public importance, or (e) Is a disposition of a case in which there is a published opinion by a lower court or administrative agency, unless the panel determines that publication is unnecessary for clarifying the panel's disposition of the case, or (f) Is a disposition of a case following a reversal or remand by the United States Supreme Court, or (g) Is accompanied by a separate concurring or dissenting expression, and the author of such separate expression requests publication of the disposition of the Court and the separate expression.

9th CIR. R. 36-2 (Rev. Jan. 1, 2012).

78. Indeed, there is some evidence that a judge on a given panel may threaten to write a dissent to a proposed opinion unless the panel agrees to issue the opinion as unpublished, thus using the threat of a dissent to push the opinion into the less salient unpublished category. *See* Mitu Gulati & Catherine McCauliff, *On* Not *Making Law*, 61 L. & CONTEMP. PROBS. 157, 204 (1998) ("Two judges inclined to reverse in a close case might agree to affirm without opinion when the third judge threatens to dissent from a published opinion ordering reversal; on the other hand, the third judge may agree to vote for an affirmance if only a nonprecedential JO [Judgment Order] is used."). If we were focusing on the prevalence of dissents, the possible suppression of dissents to avoid publication might be relevant. But there is no evidence of partisan, racial, or gender differences with respect to which judges might threaten (or might respond to a threat) to issue a dissent unless an opinion is unpublished. So there is no reason to believe that our results are systematically affected by the effects of these threats.

79. *See* Rachel Brown, Jade Ford, Sahrula Kubie, Katrin Marquez, Bennett Osdiek & Abbe R. Gluck, *Is Unpublished Unequal? An Empirical Examination of the 87% Nonpublication Rate in Federal Appeals*, 107 CORNELL L. REV. 1, 35–41 (2021) (comprehensively reviewing differences between published and unpublished opinions).

Shepard's provides further reason not to include unpublished opinions in our study. As compared to published opinions, unpublished opinions contain fewer average *Shepard's* treatments of earlier cases. Published opinions contain an average of 1.9 *Shepard's* treatments, and unpublished opinions contain an average of 0.5.[80] This is not surprising, given that the unpublished designation is for opinions that are not designed to make new law for the circuit, and *Shepard's* treatments are extensive discussions of earlier cases that shed important light on their precedential value. Put differently, the point of *Shepard's* is to describe the way in which opinions grapple with earlier cases as they help to shape the law, and the point of unpublished decisions is to have a category of opinions that are particularly straightforward and thus do not need to grapple with earlier cases. Focusing on published opinions removes 66.4% of the cases, but only 11.1% of the *Shepard's* treatments, in our dataset.

Finally, denoting an opinion as unpublished not only indicates less significance but also helps ensure that the opinion will in fact have less significance. Just as legislators are more likely to focus on more important bills, we would expect judges to emphasize published opinions. Those published opinions are, by circuit rule and court practice, the opinions on which future judges are likely to rely. We can thus refine our hypotheses above by limiting our focus to published opinions, on the theory that such opinions are the ones on which judges will actually focus.

B. Beyond Partisanship – Racial and Gender Differences

We turn now to race and gender. The underlying theory is that partisanship may not capture elements that judges find relevant in deciding which opinions to follow or cast doubt on, and that there may be commonalities among judges along race or gender lines that partisanship does not capture. But different theories yield different hypotheses.

80. These numbers count *Shepard's* treatments of U.S. Courts of Appeals opinions decided between 1974 and 2017, to make them comparable to the dataset that we use for our analysis. Many of the treating cases also contain treatments of Supreme Court opinions, district court opinions, and pre-1974 federal appellate opinions. The average number of total treatments in published opinions in our dataset is 2.7, and for unpublished opinions it is 1.6. These larger averages reflect all treatments, including those of Supreme Court opinions etc.

The studies of racial and gender differences in voting noted in Part II suggest one possibility: judges may have a particular understanding of and sensitivity to particular kinds of issues, perhaps flowing from their expertise and lived experiences.[81] A Black judge, for example, might have a deeper understanding than White judges (including within the same party) of the ways that White people engage in racial discrimination because she is more likely to have been subject to racial discrimination and to have seen multifarious forms of it. The same reasoning might apply to sex discrimination cases with respect to female versus male judges (again, including within the same party). If so, then we might expect to find differences between co-partisan judges of different races for race-salient cases and between co-partisan male and female judges for gender-salient cases.[82] Democratic Presidents have been more likely to nominate Black, Hispanic, and female judges than have Republican Presidents, so looking at racial and gender differences within party is important because it rules out the possibility that any observed differences are actually driven by differences in partisanship.

Because the theory is that judges are favorably treating opinions whose substantive approach they agree with, differences in these treatments based on race or gender can be understood as policy-motivated differences.[83] But note that such policy-specific motivations are distinct from broader, more all-encompassing ideological

81. *See supra* notes 51–54, 65 and accompanying text.

82. The terms race- and gender-salient refer to particular categories of cases that previous studies have suggested might divide judges along race or gender lines. *See supra* notes 51–58, 62–65 and accompanying text. For the list of the race-salient and gender-salient categories of cases that we compiled from previous studies and use in this Article, see *infra* Appendix A1. There are eighteen race-salient categories and twelve gender-salient categories. The race- and gender-salient categories are also included in the thirty-eight ideologically salient categories (unsurprisingly, categories that may cut along race or gender lines may also cut along ideological lines).

83. Given the overlap between race/gender and ideology, it may be that race- and gender-salient cases are the most ideological of all cases. If so, then the categories of race- and gender-salient cases would be best understood as purer measures of ideology than the broader category of ideologically salient cases. The literature has not established such a relationship among these categories, however. Instead, studies have put forward race- and gender-salient categories as likely to have particular significance along race or gender lines without indicating that they are more purely ideological. And we draw our categories of ideologically salient, race-salient, and gender-salient categories from the existing literature. So we have no basis for concluding that the race- and gender-salient categories are the most ideological of all cases.

motivations, as they arise from issues that are particularly salient with respect to the lived experience of race and gender.

A different possibility would suggest racial and gender differences in a wider range of cases than particularly race- or gender-salient cases: maybe there are broad ideological differences based on race or gender, similar to those based on partisanship, that go beyond what partisanship alone reveals. Indeed, there is some evidence that within a given party, Black and Hispanic judges differ in their political ideology from White judges, and female judges differ ideologically from male judges—with White men being the most conservative group within each party.[84] If so, race and gender might reflect some important elements of political ideology that partisanship does not capture. The existing evidence for this proposition comes from the analysis of pre-confirmation campaign contributions of judges, not judges' behavior.[85] And, as we noted in Part II, the studies addressing judges' behavior have focused on judges' votes and have generally found racial and gender differences only in a subset of race-salient and gender-salient cases. No other study has had access to our comprehensive data, and none has been able to measure racial or gender differences in substantive treatments or to measure racial or gender differences across ideologically salient cases. Our dataset, by contrast, allows us to examine opinions' reasoning for both the more specific categories of cases that studies have posited as particularly salient for race or gender purposes as well as the broader category of cases that studies have found to be ideologically salient more generally.

Insofar as there are broad ideological differences within parties based on race or gender akin to differences based on partisanship, then just as we might expect Democratic authors to differ from Republican authors in their treatments of earlier opinions across a wide range of case categories (because of ideological differences), we might expect similar differences within parties based on race and gender across a wide range of ideologically salient cases (again, because of ideological differences). So with respect to the thirty-eight ideologically salient case types we identified, we might expect

84. *See* Sen, *supra* note 15, at 394 tbl.4 (relying on preconfirmation campaign contributions to identify ideological differences within party based on race and gender).

85. *Id.*

female judges to be more likely than male co-partisans to follow majority opinions written by other female co-partisans, and for the converse to be true with respect to opinions by male judges.[86] And we might expect a similar effect based on race. Race and gender would provide information about political ideology beyond what party membership reveals about ideology.

These two possible effects are independent of each other. Maybe, for example, there is a gap between male and female judges in ideologically salient cases but a bigger gap between male and female judges in gender-salient cases, which would provide support for both types of effects. Or it could be that there is a gender difference for all ideologically salient cases but no greater difference for gender-salient cases, or conversely that there is a gender difference within gender-salient cases but not for ideological cases more generally. On the other hand, if one observed gender differences for the subset of ideologically salient cases that excludes gender-salient cases and smaller gender differences for the gender-salient cases, then this would indicate that the observed "gender" differences are less about gender and more about general ideology/political preferences. And the same possible comparisons exist for race—comparisons among all ideologically salient versus among race-salient cases would yield information about the degree to which differences were broadly ideological or more narrowly focused on areas relevant to expertise and lived experiences.

C. Solidarity Effects

The general ideological and more policy-specific hypotheses discussed in the sections above are the ones with the strongest grounding arising out of those studies and the theory underlying them. But our data cannot establish that the explanation for any party differences or intraparty racial or gender differences is general ideology or more specific differences, as opposed to something else. What else can explain party, racial, or gender differences in treatments of earlier cases? We have no other

86. As we discuss in Sections V.B and V.C, because of the lack of Black, Hispanic, and female judges in the early years of our study, our investigation of racial and gender differences looks at substantive treatments of Black-authored, Hispanic-authored, and female-authored opinions and not White-authored or male-authored opinions.

hypotheses that are grounded in studies of judicial behavior. But another possibility occurs to us as plausible: perhaps there is an in-group preference that affects behavior, which we might call a solidarity effect.[87] On this theory, a judge might choose to follow an opinion written by a judge of the same party, race, or gender not because she had a greater affinity for the substance of that judge's opinion but instead because of their shared party, race, or gender. Following the opinion would be a way of supporting the colleague and the group.

This solidarity effect differs from the broad ideological and more policy-specific explanations in an important way regarding knowledge of, and interest in, the identity of opinion authors. Insofar as any of the broad ideological or more policy-specific differences discussed above exist, it could be that the later judge is influenced by the identity of the opinion author. The later judge could use the party, race, or gender of the opinion author as a relevant factor (or even the sole factor) in identifying substantively attractive opinions. In this way, the later judge would be using party, race, or gender as a marker of ideology/policy. But note that neither the general ideological nor the more policy-specific hypotheses discussed above depend on the later judge knowing the party, race, or gender of the authoring judge. As we noted in the introduction, a later judge might follow an ideologically congenial opinion without noticing the identity of the author. Similarly, a Black/Hispanic or female judge might follow an

87. A preference for members of one's own group is often called in-group preference, in-group favoritism, or in-group bias. *See, e.g.*, Jerry Kang & Kristin Lane, *Seeing Through Colorblindness: Implicit Bias and the Law*, 58 UCLA L. REV. 465, 476 n.37 (2010) (discussing in-group preference); Robert J. Smith, Justin D. Levinson & Joë Robinson, *Implicit White Favoritism in the Criminal Justice System*, 66 ALA. L. REV. 871, 895 (2015) (discussing in-group favoritism); John T. Jost, Mahzarin R. Banaji & Brian A. Nosek, *A Decade of System Justification Theory: Accumulated Evidence of Conscious and Unconscious Bolstering of the Status Quo*, 25 POL. PSYCH. 881, 902 (2004) (same); Laurie A. Rudman & Stephanie A. Goodwin, *Gender Differences in Automatic In-Group Bias: Why Do Women Like Women More than Men Like Men?*, 87 J. PERSONALITY & SOC. PSYCH. 494, 494–95 (2004) (discussing in-group bias); Celina M. Chatman & William von Hippel, *Attributional Mediation of In-Group Bias*, 37 J. EXPERIMENTAL SOC. PSYCH. 267, 271 (2001) (same); Jennifer A. Richeson & Nalini Ambady, *Who's in Charge? Effects of Situational Roles on Automatic Gender Bias*, 44 SEX ROLES 493, 494 (2001). We use the term "solidarity effect" for two reasons. First, some of the literature on in-group preferences uses the term to focus on implicit preferences, and in this discussion we are not assuming that any preference is merely implicit. *See, e.g.*, Smith et al., *supra*. Second, the terms in-group preference, in-group favoritism, and in-group bias may have a negative connotation, and we want to avoid any such connotation.

opinion that discusses racial or sex discrimination in a sophisticated and powerful way without noticing the identity of the author. Indeed, insofar as party is correlated with ideology, race is associated with a more sophisticated understanding of race, and gender is associated with a more sophisticated understanding of sex, then we might expect party, racial, and gender differences in the substantive treatments of opinions in ideologically salient, race-salient, and gender-salient cases even in a counterfactual world in which judges' names (or other explicit identifiers of party, race, or sex) were not included in opinions: later judges would find the substance of the opinions written by those of the same party, race, or gender more attractive and thus presumably be more likely to follow them.

By contrast, the solidarity effect depends on later judges not merely noticing but in fact acting on the identity of opinion authors. The whole point of the solidarity effect is that later judges prefer to follow the opinions written by judges of the same party, race, or gender, and that of course requires that the later judges be aware of the party, race, or gender of the author.

This does not mean that the preference for following opinions by judges of the same party, race, or gender need be conscious: the later judges might have such a preference but not acknowledge it even to themselves.[88] But the solidarity effect does entail that judges in fact act on a preference to follow the opinions of those of the same party, race, or gender when opinions written by judges of a different party, race, or gender are at least as ideologically congenial to them.[89]

88. That said, the solidarity effect may be more likely to be conscious than an ideological/policy effect would be. As to the latter, there is no cognitive dissonance entailed in a judge thinking that a particular form of reasoning is powerful without taking the added step of associating that reasoning with a particular worldview. It may be naïve or betray a lack of intellectual curiosity to fail to consider how the preferred reasoning accords with a particular worldview, but it need not entail the judge hiding anything from herself. By contrast, if a judge prefers to follow the opinions of those of the same party, race, or gender because of that shared characteristic but does not acknowledge that preference to herself, she would seem to fail to understand her motivations. Whereas relating a particular form of reasoning to ideology/policy entails making a (possibly nonobvious) connection, no such connection is required in relating the identity of the opinion author to the identity of the opinion author—they are one and the same. It is of course possible that any in-group preference is unconscious (there is an extensive literature on implicit group preferences, *see supra* note 87). But it seems at least as possible that a judge who preferred to follow the opinion of someone of the same party, race, or gender would be conscious of doing so.

89. By hypothesis, judges are not choosing to follow opinions based on ideological congeniality.

The proposition that judges' decisions are influenced by the identity of earlier opinion authors may seem implausible. Circuit judges write many opinions. Is it likely that they note the identity of the opinion author of many of the opinions they follow? We cannot know, of course, but we believe it is plausible that judges are aware of the author's identity for at least some of the opinions they follow. First, the average published opinion follows 1.1 opinions and negatively treats 0.5 opinions, and the average unpublished opinion follows 0.4 opinions and negatively treats 0.1 opinions. A busy judge likely is not aware of the identity of the authors of all the opinions she cites, but she may well note the identity of the author of the one or two opinions she chooses to rely on as controlling.[90] Second, 71.2% of *Shepard's* treatments are to cases within the same circuit, the median time difference between the treating and treated opinion is 5.1 years, and the mean time difference is 7.5 years. So if later judges see the name of the opinion author, they are likely to know that author's identity.

As we noted above, the idea that ideology matters reflects an assumption that judges want to move the law in what they regard as a positive direction, and that a judge's definition of "positive" will be correlated with ideology, with the result that judges will thus want to push the law in an ideological direction.[91] The connection between identity and judicial motivations is less obvious. What is the non-ideological reason why a judge might choose to follow opinions based on shared party, race, or gender? The answer must be that one (or more) of those characteristics is important to judges, and that importance translates into influence on judicial behavior. As to importance, the idea is that some characteristics are likely to be particularly significant to the self-definition of those who share them. Race and gender, for example, are likely more central to many people's self-definition than are many other characteristics they may have (e.g., height).[92] That importance may manifest itself as influence for two related reasons. First, the importance of the characteristic may lead those who share it to feel an allegiance with one another and a desire to enhance the status of others with that characteristic. Second, enhancing the

90. On what is entailed in *Shepard's* identifying an opinion as following another opinion, see *supra* notes 30–35 and accompanying text.

91. *See supra* note 43 and accompanying text.

92. *See supra* notes 52, 54, 56, 64 & 87 and accompanying text.

status of others with a shared characteristic necessarily means enhancing one's own status.

This effect is probably fairly attenuated with respect to party. There are many judges of each party, and that large size would seem to diminish the level of solidarity and the attractiveness of trying to enhance the status of the members of one's party. For race and gender, though, the story seems more plausible. As with party, there are so many White judges that any solidarity effect among White judges would likely be very small. But there have been relatively few female, Black, and Hispanic judges, and it is possible that female, Black, and Hispanic judges might feel an allegiance to the small number of other judges who share their race or gender and be aware (consciously or unconsciously) that following the opinion of a judge who shares that characteristic enhances the status of everyone who shares it (including the judge who issues the positive treatment).

The discussion above leads to two possible forms of a solidarity effect. One is that judges have a fairly consistent general preference for enhancing the status of those of the same race or gender for non-ideological reasons. When considering which of several similar opinions to follow, judges will be inclined to follow the opinions of those of the same race or gender. If, say, female judges prefer to follow the opinions of other female judges, then we should expect to see intraparty gender-based differences in positive treatments as soon as there were enough female judges to allow for meaningful comparison that continue throughout the remainder of our study period.

A different possibility is that female, Black, and Hispanic judges felt a level of solidarity and kinship arising from their shared status as a small group of relative trailblazers. Solidarity arising from such a trailblazer effect would suggest a pattern to treatment differences based on race and gender: when there were very few female, Black, or Hispanic judges, we might expect greater differences between female and male judges, and between Black or Hispanic and White judges, on the theory that the benefits of positive treatments would be particularly meaningful for the first few judges with a particular characteristic.

Once there was a critical mass of female, Black, or Hispanic judges, we might expect those differences to be reduced, on the theory that, with a critical mass of judges with a particular attribute,

the sense of being a tiny cohort of trailblazers would no longer be as powerful.[93] With larger numbers, there might still be some solidarity effect, but it might be smaller because the relevant group would perceive itself to be more established and less in need of proving itself. For some (or all) of these groups, we might not yet have reached that point of critical mass, and current female, Black, and/or Hispanic judges might see themselves as trailblazers. If so, then we might expect to see a consistent difference in intraparty substantive treatments based on race or gender, rather than a decline in the most recent time period.

A trailblazer effect would thus suggest one of two possibilities. We might see larger gender differences in *Shepard's* treatments in the middle two time periods of our study (1985–95 and 1996–2006), on the assumption that there were too few female judges until 1985 and a sufficient number by the last time period in our study (2007–17) that female judges would be less likely to see themselves as trailblazing members of a tiny cohort. For Black judges, the trailblazer effect might have extended longer, and for Hispanic judges longer still.[94] Or, as suggested above, perhaps for some categories the trailblazer effect would persist, such that we might see a consistent effect with respect to race or gender in the last three time periods of our data (1985–2017).

D. Measures and Data Limitations with Respect to Race and Gender

We have identified three possible mechanisms that would give rise to racial and gender differences: racial or gender differences specific to issues about which Black, Hispanic, or female judges have particular expertise; racial or gender differences reflecting broader ideological differences, akin to those separating Democrats and Republicans; and racial or gender differences reflecting solidarity effects. We test all three by examining different categories

93. Laura Moyer and Susan Haire found such a trailblazer effect that dissipated over time. Specifically, they found differences in female and male judges' likelihood of voting in favor of plaintiffs in sex discrimination cases, but only for the earliest cohorts of judges. *See* Moyer & Haire, *supra* note 56.

94. There were at least two female judges in most circuits by 1992 (and all but two circuits by 1998). By contrast, in our time period three circuits never had more than one Black judge sitting at any given time, and it was not until 2002 that a majority of circuits had at least two Black judges. And the numbers for Hispanic judges are bleaker: in our time period, four circuits never had a Hispanic judge, and only four had more than one. *See infra* Appendix A2.

of cases, breaking them up into different time periods, and looking at differences within parties.

Specifically, we examine whether there are gender differences in treatments within party. We do this by subsetting the treating cases to just those written by the members of a given party and then looking to see whether female co-partisan opinion authors are more likely than male co-partisan opinion authors to follow opinions written by female co-partisans. We then do the same analysis focusing on gender-salient cases, more broadly ideologically salient cases, and the group of ideologically salient cases that excludes gender-salient cases. And we divide our forty-four years of data into four time periods to examine whether there are changes over time.

Similarly, we investigate whether the race of judges is associated with positive treatments of opinions written by same-race co-partisan judges—that is, holding partisanship constant.[95] We use biographical data from the Federal Judicial Center to label judges who self-identify as Black, non-Hispanic White, or Hispanic. And we do the analogous analysis focusing on race-salient cases, ideologically salient cases, and ideologically salient cases minus race-salient cases, and we divide our data into four time periods.[96]

Analysis of racial and gender differences depends on there being a sizable pool of cases written by Black, Hispanic, and female judges as well as at least one Black, Hispanic, or female judge in the circuit who can follow one of those earlier cases. And given the strong tendency of judges to follow cases within their circuit, the most substantively relevant results will arise when there are at least two Black, Hispanic, or female judges in a given circuit, such that all judges can choose to follow an opinion they did not write that was written by someone of the same or a different race or gender.

As section III.C indicated, this is a modest data limitation with respect to gender but a significant one with respect to race.[97] We could avoid these data limitations if we combined all circuit judges together into a single group undifferentiated by circuit or year, as

95. We provide a full explanation of these intraparty racial and gender measures in Sections V.B and V.C.

96. As with all the analysis in this Article, we estimate whether judges in each group are more likely to follow opinions from judges of their own group than are judges from the other group(s) after adjusting for circuit, year, and circuit-year combinations. *See supra* text accompanying note 70.

97. *See supra* note 94.

then there would be more than enough Black and Hispanic judges to allow for comparisons. We do not perform such an analysis because such an agglomeration of judges runs the risk of inaccurate findings. Because 71.2% of *Shepard's* treatments are to cases within a given circuit, it would be problematic for us to treat all judges or cases—even within a given year—as an undifferentiated whole. So just as we condition on each circuit-year combination when looking at partisanship, we do the same here.

With the advantage of avoiding spurious results comes the disadvantage of inferences that are only relevant for a narrowly defined population of cases from less than all circuit-years. The problem is most acute with respect to race. Our focus on circuit-year combinations means that for most such combinations there will not be enough Black or Hispanic judges in prior years to have many opportunities for substantive treatments of opinions written by minority judges. This is a data limitation that we have no control over. Importantly, it implies that our results—particularly those regarding White-Hispanic comparisons within Republican appointees—are only representative of a narrow set of cases from a limited set of circuits and years.

IV. DATA AND RESEARCH DESIGN

A. Data

To construct our data, we gathered all published and unpublished federal appellate opinions in the Lexis database issued between 1974 and 2017.[98] We separately identified each substantive *Shepard's*

98. We received the data directly from Lexis, which sent us all the circuit court opinions (including opinions issued by a single judge or two judges) in its database for our time period. We spent nine months analyzing the data for any possible lacunae or discrepancies and found none. We also compared the published cases from Lexis with the cases available from the Caselaw Access Project, https://case.law, and found greater than 99.9% agreement. We removed the opinions issued by a single judge or only two judges, which constituted approximately 1.4% of the observations, as we found that most of them were not decisions on the merits.

All the available online sources of opinions (including Lexis) fail to include some unpublished orders and opinions. *See* Michael Kagan, Rebecca Gill & Fatma Marouf, *Invisible Adjudication in the U.S. Courts of Appeals*, 106 GEO. L.J. 683, 685 (2018) (finding that some unpublished decisions in deportation proceedings are not in online databases, although Lexis has the best coverage); Merritt E. McAlister, *Missing Decisions*, 169 U. PA. L. REV. 1101, 1103 (2021) (finding that some unpublished orders and opinions are not in online databases). Apparently, the various circuits consider some unpublished orders and opinions to be

treatment in every court of appeals opinion. We obtained the data for court opinions, judges, and *Shepard's* treatments from LexisNexis. We focus on substantive treatments within majority opinions generated by panels of the federal courts of appeals.[99] Thus our measure of *Shepard's* treatments does not include treatments of district court opinions or Supreme Court opinions. In addition, we exclude the Federal Circuit from our analysis, because it has relatively few of the ideologically, race-, and gender-salient cases that we want to measure.[100] The full dataset comprises 670,784 federal appellate majority opinions and 648,226 *Shepard's* treatments.

In our data, 69.6% (451,289) of the treatments are positive, 8.5% (55,169) are neutral (which, as we noted previously, we drop precisely because they are neutral), and the remaining 21.8% (141,768) are negative.[101] As we noted in section III.C, the average published opinion has 1.1 positive treatments and 0.5 negative treatments of earlier opinions. More than 70% of the treatments address opinions decided within 10 years of the original opinion. Only 7.4% of the treatments in our data show a gap of more than 20 years from the original opinion, so it is relatively uncommon for an opinion to rely on opinions from an earlier era.

Lexis provided names of the authoring judge of an opinion for 301,337 cases, which amounts to 44.9% of our data. Most of these

sufficiently trivial that they are not even included among the unpublished orders and opinions that are sent to the online databases. As Kagan, Gill & Marouf note, the fact that these opinions are not available in online databases means that these opinions are not merely nonprecedential but also invisible. Kagan, Gill & Marouf, *supra*, at 689 (noting two categories of "invisible" decisions—"Nonprecedent, invisible decisions" and "Nonmerits decisions (invisible)"). Indeed, these opinions would be invisible not only to lawyers but also to judges and clerks, except for those few who might have worked on one of the invisible cases. *See* McAlister, *supra*, at 1149. The invisible opinions are thus not cited (much less discussed), which of course continues their invisibility. This point is significant for our purposes because our focus in this Article is on courts' treatments of earlier opinions. Invisible opinions, by being invisible, are not available for later treatment and thus fall out of the denominator. That said, if there were patterns of invisible opinions that related to the party, race, or gender of the judges deciding the cases, that could bias our results. But there is no reason to believe that any such patterns exist. *See* McAlister, *supra*, at 1146–47; *supra* notes 76–80 and accompanying text.

99. As we note in Part I, *Shepard's* does not code dissenting or concurring opinions as positive or negative (for good reasons), so our data encompass majority opinions. *See supra* note 28.

100. *See supra* note 4.

101. As we noted in Part I, "followed" constitutes more than 99.99% of the positive treatments, so we refer to following and positive treatments interchangeably. *See supra* note 27 and accompanying text.

cases are published cases. We then used the data from the Caselaw Access Project to cross-reference and supplement the author names.[102] We dropped from our analyses 39,906 cases with author names indicating that they were per curiam opinions.[103]

We obtained data on judge-level characteristics such as judges' birth year, commission year, race, and gender from the Federal Judicial Center. Similarly, we used data from the Federal Judicial Center to identify the President who most recently nominated each judge.

Lexis assigns multiple topic headers to each case to identify the legal issue areas that a given case addresses. An average opinion has ten topic headers. This reflects the specificity of topic headers—an ordinary case does not cover ten completely different areas of law, but it might cover ten closely related and highly specific topics. Each topic header is a hierarchy moving from broad to more specific categories of law. A typical topic header is "Labor & Employment Law>Discrimination>Gender & Sex Discrimination> Evidence>Burdens of Proof>Burden Shifting."[104]

We categorized each opinion based on whether at least one of its topic headers contains the topic-identifying keyword for one of the thirty-eight ideologically salient issue areas, e.g., Search and Seizure, Immigration Law, and Sex Discrimination.[105] In our data, 213,619 cases are assigned topic headers that pertain to the thirty-eight ideologically salient issue areas we identified. We refer to this group of cases as the ideologically salient subset.

In each empirical analysis that follows, we fit the same model specifications to three different subsets—unpublished and

102. *Caselaw Access Project*, LIBR. INNOVATION LAB, https://case.law (last visited Sept. 16, 2023). The Caselaw Access Project provides open access to raw texts of all published opinions in U.S. courts. This allows us to compare the names of authoring judges of the published opinions in Lexis and Caselaw. We matched cases in the two datasets based on the case title and decision date. After multiple steps to adjust for different formats and styles, we supplemented the authoring judge for 9,901 of the opinions we received from Lexis with data from the Caselaw Access Project.

103. For some cases, the data shows the author's name as "Per Curiam," and for others the data indicates that the case is per curiam but shows the names of all participating judges. We drop these cases and keep only opinions that show one judge name as the author.

104. This is a topic header from Kidd v. Mando Am. Corp., 731 F.3d 1196 (11th Cir. 2013), a case we randomly selected from our dataset for purposes of illustrating the topic headers.

105. The thirty-eight ideologically salient issue areas are in Appendix A1. The topic categories are not mutually exclusive under our coding rule. In other words, an opinion can, and likely does, have multiple topic categories.

published combined, published only, and published and ideologically salient. There are 225,465 published opinions and 106,804 published and ideologically salient opinions in our data. In addition, some analyses make use of even more fine-grained subsets of published and race-salient cases (of which there are 70,066) and published and gender-salient cases (of which there are 46,458).

B. Research Design

Because we have collected essentially all opinions issued by federal courts of appeals, we effectively have the entire population of data from the time period in question. We thus do not need to rely on random sampling of cases or other methods to ensure that our descriptive claims are accurate.

It is also worth reiterating that our decision to use *Shepard's* treatments as our primary outcome variable has several benefits. Our results do not depend on our own substantive judgments about how to code outcomes, and past research has shown the *Shepard's* measures to be reliable and valid.[106]

Further, we expect that our approach of looking at how sitting judges make use of the opinions of past judges is much more likely to produce measures that are comparable over time than approaches that focus on which litigant prevailed in a dispute and/or that attempt to discern the ideological valence of a decision. Measures that rely on the identity of the prevailing side run afoul of all manner of serious and not-so-serious selection issues, as strategic litigants will condition their litigation strategy on their expectation of prevailing on the merits.[107] The win rates of certain types of litigants are not of interest to us in this Article. We are also not interested in what would happen as a result of counterfactual changes in litigation strategies. In short, the behavior of litigants—along with the concomitant selection issues—is not of concern here.

Measures that require researchers to make substantive judgments about what the "liberal" or "conservative" outcome is within certain types of cases can miss key aspects of the legal reasoning and can also be difficult to compare over time. The meaning of "liberal" and "conservative" is timebound, and the types of disputes clearly change

106. *See supra* notes 30–36 and accompanying text.

107. *See* George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1, 11 (1984).

over time.[108] Because judges always have and will seek to justify their decisions by relying on earlier decisions, there is little question of the comparability of our outcome measures over time. Further, because we are interested in the body of law as it exists and not the body of law that would counterfactually exist had some disputes not settled (and we are not interested in the win rates of certain types of litigants), we do not need to be concerned about the litigation strategies of the parties.

As we noted in Part III, when comparing the behavior of Democratic appointees to Republican appointees, White judges to Black judges, White judges to Hispanic judges, and female judges to male judges, we attempt to get as close as we can to apples-to-apples comparisons. We do this by adjusting for circuit, year, and circuit-year fixed effects. The exact form of this adjustment is described in more detail in section C below. Further, we also report time-period specific differences to better assess the extent to which behavior is changing over time.

C. Estimation and Inference

To adjust for circuit, year, and circuit-year effects, we estimate and report what are known as average controlled differences using overlap weights.[109] Put simply, for a particular comparison, say Democrat versus Republican, this approach works by first estimating the probability, within each circuit and year, that each case is authored by a Democrat and the probability that each case is authored by a Republican.[110] The behavioral difference of interest, say the difference in positive treatments of Democratic-authored opinions, is then defined as a weighted average difference where, in our running example, the weights would be proportional to the probability the case was authored by a Democratic appointee times the probability that the case was authored by a Republican appointee. In this running example, this average controlled difference is estimated by weighting the Democratic-authored opinions by the probability that they could have been authored by

108. *See* Carolyn Shapiro, *Coding Complexity: Bringing Law to the Empirical Analysis of the Supreme Court*, 60 HASTINGS L.J. 477, 480 (2009); *supra* notes 18–21 and accompanying text.

109. *See* Li, Zaslavsky & Landrum, *supra* note 70; Fan Li, Kari Lock Morgan & Alan M. Zaslavsky, *Balancing Covariates via Propensity Score Weighting*, 113 J. AM. STAT. ASS'N 390, 391 (2018).

110. These probabilities are referred to as propensity scores in the literature. Li, Morgan & Zaslavsky, *supra* note 109.

a Republican and weighting the Republican-authored opinions by the probability they could have been authored by a Democrat.[111]

As discussed by Fan Li, Kari Morgan, and Alan Zaslavsky, this approach to weighting the data serves to perfectly balance the measured covariate distributions across the comparison groups.[112] In other words, if we are looking at a comparison of Democratic judges to Republican judges, the weighted fraction of Democratic-authored opinions from a particular year and circuit will be the same as the weighted fraction of Republican-authored opinions from that same year and circuit. This is highly desirable, as it eliminates circuit-, year-, and circuit-year-specific factors as confounding variables.

Further, since the overlap weights are proportional to a probability (say the probability of a case being authored by a Democrat) times one minus that probability, the cases that will get the most weight are those with a 50-50 chance of being decided by either type of author in the comparison. Not only is this part and parcel with producing covariate balance, but it also importantly downweights cases from circuit-years where the comparisons of interest are simply difficult if not impossible to make in a credible fashion. For instance, if there are no Hispanic judges in a particular circuit and year, then it would not be meaningful to make White-Hispanic comparisons within that circuit and year. Our estimation approach automatically gives the substantive treatment decisions from that circuit and year zero weight.

We construct standard errors and confidence intervals using the nonparametric bootstrap.[113]

Finally, it is important to note that we estimate a large number of average controlled differences in this Article. A concern when conducting many hypothesis tests (or equivalently looking to see whether many *p*-values fall below a threshold) is the high likelihood of making many false discoveries (i.e., incorrect rejections of true null hypotheses). We guard against this by employing the methods

111. *See id.*

112. *See id.*

113. *See* BRADLEY EFRON & ROBERT J. TIBSHIRANI, AN INTRODUCTION TO THE BOOTSTRAP 42–45 (1993) (explaining the calculation of nonparametric bootstrap standard errors and confidence intervals).

of Benjamini and Yekutieli to control the false discovery rate.[114] We also use the associated method of Wright to adjust the *p*-values.[115]

V. PRIMARY RESULTS

Having laid out our research questions, research design, and strategy for estimation and inference, we turn to the results of our study. In all these analyses, we attempt to minimize imbalances in the data by focusing on average controlled differences using overlap weights and adjusting for circuit, year, and circuit-year indicators.[116] As we noted in Part IV, this estimation strategy generates meaningful, apples-to-apples comparisons between the contrasting groups of interest (Democrat-Republican, Black-White, Hispanic-White, and female-male). It does so by downweighting data from circuit-years that are heavily skewed to one group—say, towards White judges and away from Hispanic judges. We begin with partisanship and then proceed to race and gender.

A. Average Controlled Differences by Author Partisanship

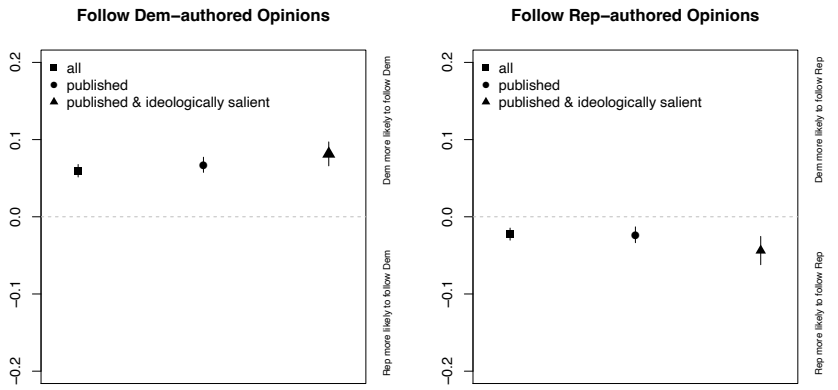
We begin by examining the role played by partisanship, which we hypothesize to play a major role in structuring judicial behavior. We start by pooling our data over the full time span of the study (1974–2017) and looking at partisan differences in the propensity to follow Democratic-authored opinions and Republican-authored opinions. The results are broken down for all cases, published cases, and published and ideologically salient cases.

114. *See generally* Yoav Benjamini & Daniel Yekutieli, *The Control of the False Discovery Rate in Multiple Testing Under Dependency*, 29 ANNALS STAT. 1165, 1166 (2001) (providing a procedure to control the false discovery rate); S. Paul Wright, *Adjusted P-Values for Simultaneous Inference*, 48 BIOMETRICS 1005, 1007 (1992) (providing a correspondence between methods to control the false discovery rate and adjustment of *p*-values).

115. *See* Wright, *supra* note 114.

116. *See supra* note 109 and accompanying text.

Figure 1

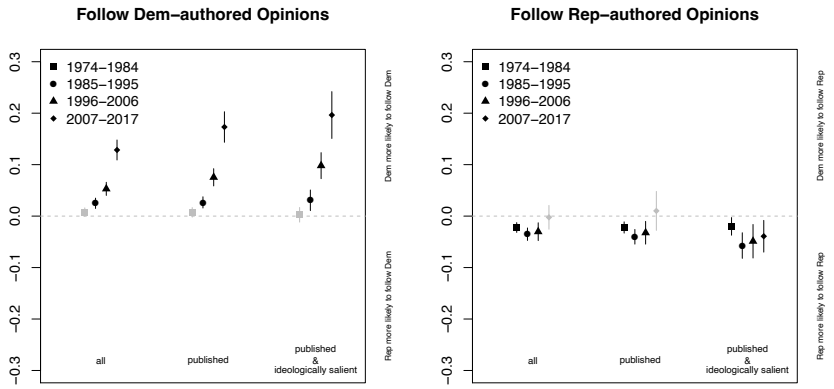


Partisan differences in positive treatments of opinions authored by a Democratic or Republican appointee for all, published, and published and ideologically salient cases. Points on the plot represent the difference in the propensity of Democratic authors versus Republican authors to follow opinions by other Democratic (left panel) or Republican (right panel) authors. Points above the horizontal line at 0 indicate a greater propensity of Democratic appointees than Republican appointees to follow the partisan-authored opinions in question. Points below the horizontal line at 0 indicate a greater propensity of Republican appointees than Democratic appointees to follow the partisan-authored opinions in question. Each panel plots three different average controlled difference estimates: the left point is for all cases (published and unpublished combined), the center point is for published cases, and the right point is for published and ideologically salient cases. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are depicted in black, and those that are statistically insignificant at the 0.05 level after adjusting for multiple testing are depicted in gray.

Figure 1 displays these results. We find statistically significant partisan differences across all cases, published cases, and published ideologically salient cases. These differences are all in the expected direction. The left panel of *Figure 1* shows that Democratic opinion authors are more likely to follow opinions authored by other Democratic appointees than are Republican opinion authors. Similarly, we see in the right panel of *Figure 1* that Republican opinion authors are more likely than Democratic authors to follow opinions written by Republican appointees. For both plots in *Figure 1* the partisan differences grow slightly stronger as we narrow our analyses down to published opinions and published and ideologically salient opinions. The partisan differences in positive treatments to Democratic-authored opinions are larger in absolute

value than the partisan differences in positive treatments to Republican-authored opinions. Further, the estimated differences are large enough to be substantively meaningful. For instance, the Democrat-Republican difference in the average positive treatments to Democratic-authored opinions is approximately 0.1, which corresponds to a Democrat giving one more positive treatment to a previous Democratic opinion than was given by a Republican in every 10 opinions.

Figure 2



Partisan differences in positive treatments of opinions authored by a Democratic or Republican appointee over time for all, published, and published and ideologically salient cases. Points on the plot represent the difference in the propensity of Democratic authors versus Republican authors to follow opinions by other Democratic (left panel) or Republican (right panel) authors. Points above the horizontal line at 0 indicate a greater propensity of Democratic appointees than Republican appointees to follow the partisan-authored opinions in question. Points below the horizontal line at 0 indicate a greater propensity of Republican appointees than Democratic appointees to follow the partisan-authored opinions in question. The shape of the points denotes the average controlled difference for different time periods, with each period encompassing 11 years. For both panels, the left four points display the average controlled differences for all cases (published and unpublished combined), the center four points show the average controlled differences for published cases, and the right four points show the average controlled differences for published and ideologically salient cases. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are depicted in black, and those that are statistically insignificant at the 0.05 level after adjusting for multiple testing are depicted in gray.

We next break these partisan author analyses down by four time periods (1974–1984, 1985–1995, 1996–2006, 2007–2017). *Figure 2* plots partisan differences in the propensity to follow opinions

written by Democratic and Republican authors. Looking at *Figure 2*, we see that Democratic opinion authors are much more likely than Republican opinion authors to follow earlier Democratic-authored opinions. Further, these partisan differences grow dramatically larger over time with the size of the differences accelerating rapidly. The differences are largest, with the greatest acceleration, within the subset of published and ideologically salient opinions. In the most recent time period, Democrats make more than 1 additional positive treatment to a Democratic-authored opinion than do Republicans in every 5 opinions, as is seen in the estimated difference being greater than 0.2. There is some evidence that Republican opinion authors are more likely than Democratic opinion authors to follow opinions written by Republicans. While these differences are significantly different from zero in the first three time periods for all cases and published cases, and significantly different from zero in all four periods for published and ideologically salient cases, they are much smaller than the corresponding differences with respect to positive treatments of Democratic-authored opinions. Further, the differences do not become larger over time, in contrast to the positive treatments to Democratic-authored opinions.

B. Average Controlled Differences by Author Race

In this section we present results on the extent to which there are racial differences in substantive treatments. The history of Black and Hispanic representation on the federal courts of appeals gives rise to two related empirical patterns that need to be dealt with when conducting this analysis. First, Black and Hispanic judges appeared on the bench in substantial numbers only in the mid-1990s and after—over twenty years after the first year of our data.[117] Second, during the 1974–2017 time period Black and Hispanic judges were more likely to be appointed by a Democratic President than a Republican President. Both facts have implications for which outcome variables are most meaningful to study.

The fact that few Black and Hispanic judges were on the bench prior to the 1980s means that the vast majority of opinions written prior to the 1980s were written by White judges. It is thus not particularly useful to look at the propensity of White and minority

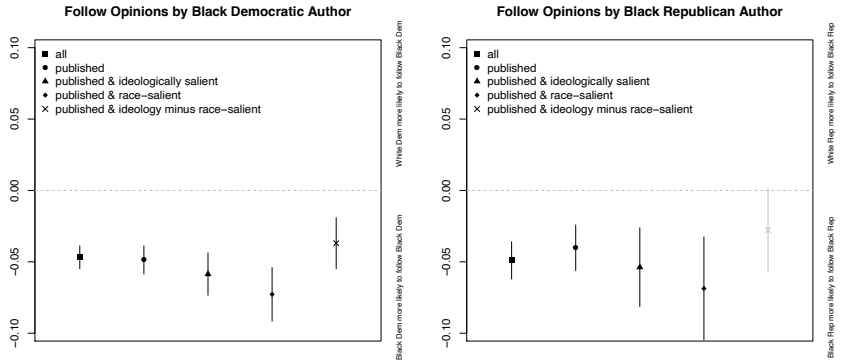
117. *See infra* Appendix A2.

judges to follow White-authored opinions without some form of adjustment, since both White and minority judges effectively had only White-authored opinions to treat until a sufficient body of minority-authored case law was developed well after the start of our data.

Relatedly, the fact that most Black and Hispanic judges—particularly in the early years of our study—were Democratic appointees means that there is also a substantial partisan skew to the minority-authored opinions that do exist. Failing to adjust for this partisan difference in the stock of minority-authored opinions that can be followed will also result in unreliable inferences about the role of race in structuring how judges treat earlier opinions.

The approach we take to deal with both issues is to condition our analysis on partisanship and to examine the extent to which White and minority judges from a given party positively treat opinions written by minority co-partisan judges. More specifically, we subset the data down to cases with substantive treatments authored by a given party and then estimate the average controlled difference between how White and either Black or Hispanic co-partisans positively treat past opinions written by Black or Hispanic co-partisans. Using positive treatments of minority-authored opinions as the outcome variable automatically adjusts for the later arrival of substantial numbers of minority judges in our data since both White and minority judges will have equal opportunity to positively treat minority-authored opinions as long as there are some minority judges (and thus some minority-authored opinions). Doing this within party adjusts for the partisan skew of the minority-authored opinions.

Figure 3

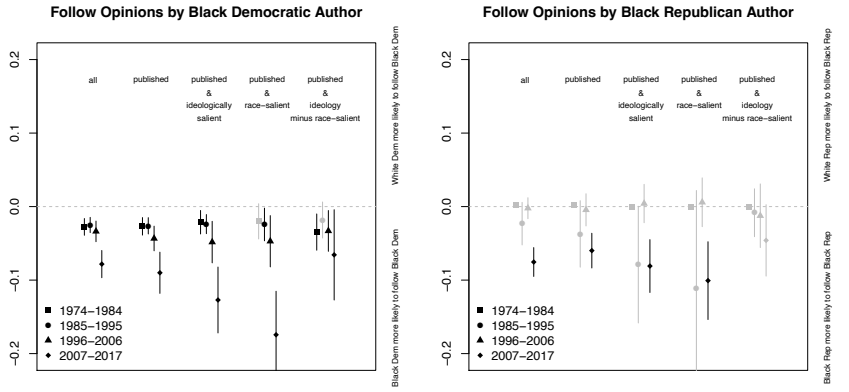


Intraparty racial differences in the positive treatments of opinions authored by a Black judge of the same party for all, published, published and ideologically salient, published and race-salient, and published and ideologically salient minus race-salient cases. Points on the left panel represent the difference in the propensity of White Democratic authors versus Black Democratic authors to follow opinions by Black Democratic authors. Points on the right panel represent the difference in the propensity of White Republican authors versus Black Republican authors to follow opinions by Black Republican authors. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are depicted in black, and those that are statistically insignificant at the 0.05 level after adjusting for multiple testing are depicted in gray.

Figure 3 displays the average controlled White-Black intraparty differences (again controlling for year, circuit, and circuit-year interactions) aggregated over the full time span of our study but broken down by case type (all, published, published and ideologically salient, published and race-salient, and published and ideologically salient excluding race-salient).[118] As expected, within party, Black judges are more likely to follow the opinions of fellow Black judges than are White judges. These differences get somewhat larger as one moves from all cases to ideologically salient and race-salient cases and diminishes somewhat for the category of ideologically salient but not race-salient. That said, the differences are generally similar in magnitude across the various subsets of cases. We discuss the implications of this in section VI.B. Further, the size of these differences is similar across parties.

118. Recall that the race-salient cases (and the gender-salient categories) are subsets of the ideologically salient categories. *See supra* note 82.

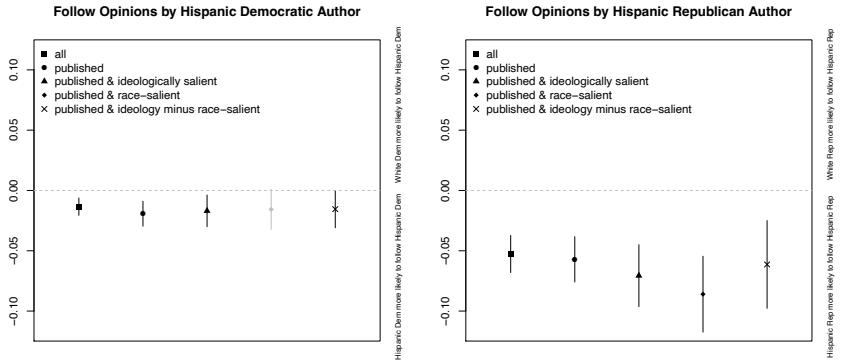
Figure 4



Intraparty racial differences in the positive treatments of opinions authored by a Black judge of the same party over time for all, published, published and ideologically salient, published and race-salient, and published and ideologically salient minus race-salient cases. Points on the left panel represent the difference in the propensity of White Democratic authors versus Black Democratic authors to follow opinions by Black Democratic authors. Points on the right panel represent the difference in the propensity of White Republican authors versus Black Republican authors to follow opinions by Black Republican authors. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are depicted in black, and those that are statistically insignificant at the 0.05 level after adjusting for multiple testing are depicted in gray.

Figure 4 disaggregates these White-Black intraparty differences by time period. Interestingly, the White-Black differences increase substantially over time. The within-Democratic White-Black differences are significantly different from zero in all but two case-subset-time-periods, but more notably the size of the difference generally accelerates rapidly over time, with the difference in the 2007–2017 time period several times larger than for the earlier periods. The size of these 2007–2017 differences is large and substantively meaningful. For instance, within the subset of published and race-salient cases, Black Democrats gave one more positive treatment to opinions written by Black Democrats in every five opinions than did their White Democratic colleagues. The 2007–2017 differences are also the largest for Republicans, although here the pre-2007 differences are not significantly different from zero.

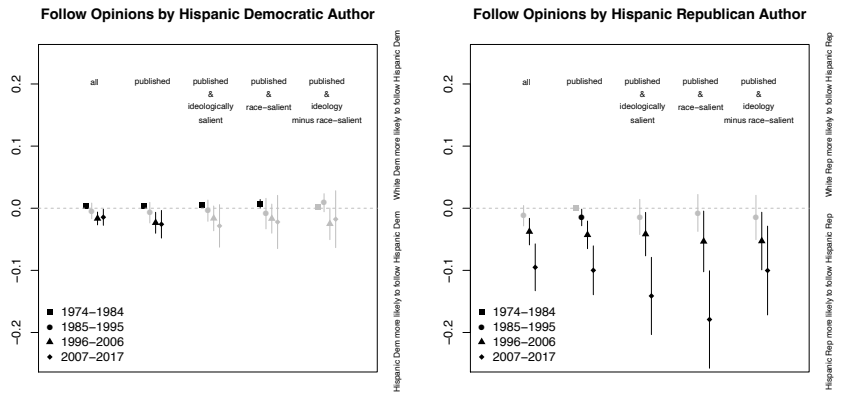
Figure 5



Intraparty racial differences in the positive treatments of opinions authored by a Hispanic judge of the same party, for all, published, published and ideologically salient, published and race-salient, and published and ideologically salient minus race-salient cases. Points on the left panel represent the difference in the propensity of White Democratic authors versus Hispanic Democratic authors to follow opinions by Hispanic Democratic authors. Points on the right panel represent the difference in the propensity of White Republican authors versus Hispanic Republican authors to follow opinions by Hispanic Republican authors. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are depicted in black, and those that are statistically insignificant at the 0.05 level after adjusting for multiple testing are depicted in gray.

In *Figure 5*, we display the White-Hispanic intraparty average controlled differences (again controlling for year, circuit, and circuit-year interactions) aggregated over the full time span of our study and broken down by case type. All but one White-Hispanic difference is significantly different from 0 and in the expected direction—Hispanic authors being more likely than White authors to positively treat past opinions written by Hispanic co-partisans. The sole statistically insignificant difference is the partisan White-Hispanic difference within published and race-salient cases among Democratic appointees. Within party, the differences across different subsets of cases are not statistically distinguishable from each other. We return to this point and discuss its substantive interpretation in section VI.B. Interestingly, the White-Hispanic differences are larger within Republican judges than within Democratic judges.

Figure 6



Intraparty racial differences in the positive treatments of opinions authored by a Hispanic judge of the same party over time for all, published, published and ideologically salient, published and race-salient, and published and ideologically salient minus race-salient cases. Points on the left panel represent the difference in the propensity of White Democratic authors versus Hispanic Democratic authors to follow opinions by Hispanic Democratic authors. Points on the right panel represent the difference in the propensity of White Republican authors versus Hispanic Republican authors to follow opinions by Hispanic Republican authors. Note that the data points for the first period in the right panel are not plotted due to a lack of data (insufficient Hispanic Republican appointees in 1974–1984). Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are depicted in black, and those that are statistically insignificant at the 0.05 level after adjusting for multiple testing are depicted in gray.

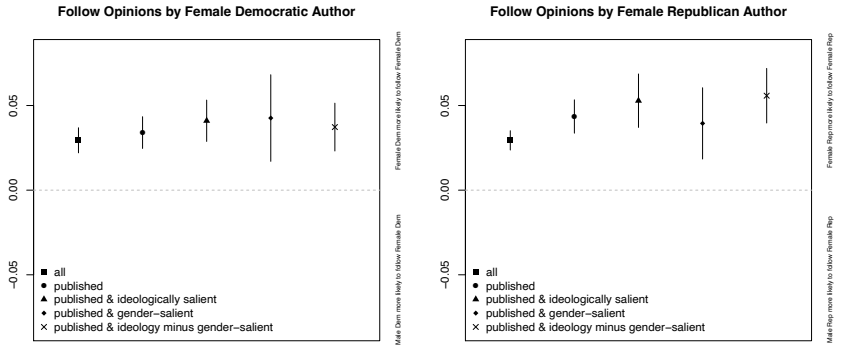
Figure 6 breaks down the intraparty White-Hispanic differences by time period. Again, among Democrats, we see White-Hispanic differences that are close to zero, with most not statistically distinguishable from zero. In the first time period, the differences for published cases, published and ideologically salient cases, and published and race-salient cases are significantly different from zero but extremely close to zero. This is driven by the very small number of opinions by Hispanic judges in this time period. The available pool of opinions in the 1974–1984 period that a Hispanic judge could follow were overwhelmingly authored by White judges. There is perhaps some slight evidence that the differences are increasing over time for this group, but again, in the subset of cases where we would expect to see the largest differences—race-salient and ideologically salient cases—the White-Hispanic

differences are not statistically significant (except for the 1974–1984 period, which is an artifact of the tiny number of cases with Hispanic judges in that time period). The White-Hispanic average controlled differences are much larger and grow much more rapidly over time for Republican appointees. Within this group, the White-Hispanic differences are either undefined or indistinguishable from zero in the first two time periods, but after 2007, we see that Hispanic Republicans are much more likely than White Republicans to positively treat Hispanic Republicans and, indeed, these differences accelerate over time.

C. Average Controlled Differences by Author Gender

The relatively recent and small female representation on the federal courts of appeals gives rise to the same sorts of issues discussed in section V.B, and we deal with this issue in the same way: we subset the data down to cases with substantive treatments authored by members of a given party and then estimate the average controlled difference between how female and male judges positively treat past opinions written by female co-partisan judges.

Figure 7



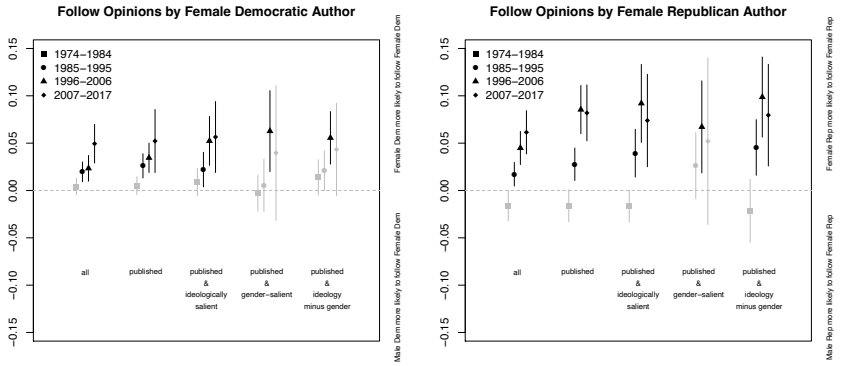
Intraparty gender differences in the positive treatments of opinions authored by a female judge of the same party for all, published, published and ideologically salient, published and gender-salient, and published and ideologically salient minus gender-salient cases. Points on the left panel represent the difference in the propensity of female Democratic authors versus male Democratic authors to follow opinions by female Democratic authors. Points on the left panel represent the difference in the propensity of female Republican authors versus male Republican authors to follow opinions by female Republican authors. Differences that are statistically significant at the 0.05 level after adjusting for multiple

testing are depicted in black, and those that are statistically insignificant at the 0.05 level after adjusting for multiple testing are depicted in gray.

Figure 7 displays these estimated average controlled differences (again controlling for year, circuit, and circuit-year interactions) aggregated over the full time span of our study. As with the race results in section V.B, we break them down by case type. We focus on gender-salient rather than race-salient cases for the obvious reason that the literature on gender differences based on expertise and lived experiences focuses on gender-salient cases. And, analogous to race, we consider the ideologically salient cases that are not gender salient.

Looking at *Figure 7*, we see that after adjusting for partisanship, gender plays a role in guiding treatment practices—with female judges being more likely than male judges to positively treat the work of female co-partisans. The size of these differences is similar for both Democratic and Republican judges—within the subset of published and gender-salient cases, slightly less than one more positive treatment from a female judge to another female judge's opinion than from a male judge in every 20 opinions. Interestingly, the average controlled differences within the published and gender-salient subset of cases are not larger than the average controlled differences within the published and ideologically salient subset. Indeed, as with the Black and Hispanic results in section V.B, none of the intraparty differences corresponding to different subsets of cases are statistically distinguishable from each other.

Figure 8



Intraparty gender differences in the positive treatments of opinions authored by a female judge of the same party over time, for all, published, published and ideologically salient, published and gender-salient, and published and ideologically salient minus gender-salient cases. Points on the left panel represent the difference in the propensity of female Democratic authors versus male Democratic authors to follow opinions by female Democratic authors. Points on the left panel represent the difference in the propensity of female Republican authors versus male Republican authors to follow opinions by female Republican authors. The shape of the points denotes the averaged controlled difference for different time periods with each period encompassing 11 years. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are depicted in black, and those that are statistically insignificant at the 0.05 level after adjusting for multiple testing are depicted in gray.

Figure 8 breaks down the average controlled differences in *Figure 7* by time period. In each case, we see evidence that the intraparty gender differences are growing larger over time. This is most apparent within the Republican subset—particularly within the subset of published and ideologically salient cases. The relatively small number of published and gender-salient cases within each of the party-time-period combinations produces a great deal of estimation uncertainty which manifests as wide confidence intervals. This level of statistical uncertainty makes it difficult to draw strong conclusions about change over time within this subset of cases. Nonetheless, we do see strong evidence of increasing gender differences over time within all cases and published cases. Further, the aggregate results presented in *Figure 7* provide strong evidence of intraparty gender differences over the entirety of our study (1974–2017).

VI. DISCUSSION

We have examined judges' substantive treatments of opinions to determine if there are substantively meaningful and statistically significant differences in those treatments based on the partisanship, race, and gender of the judges. Our findings are more nuanced and interesting than we anticipated. Pooling the data over the full span of our study (1974–2017), we find statistically significant differences with respect to positive treatments (we do not find substantively meaningful differences for the less common negative treatments). But the real story arises from looking at changes in the partisan, race, and gender differences over time. Most of the differences we find are fairly small in the early periods of our study and rise dramatically over time, becoming large and therefore substantively meaningful. Further, most of the differences are greatest for the ideologically charged categories of cases.

A. Partisan Differences

As we noted above, treatment of earlier cases is a central element of an opinion's reasoning. And, relatedly, following a case helps to increase its importance. If judges were consistently inclined to act in a partisan manner, we would expect them to treat earlier opinions by the members of their party better than earlier opinions by members of the opposite party.

Looking at all time periods combined, we find statistically significant differences in how Democratic appointees versus Republican appointees treat earlier opinions, with somewhat greater partisan differences for earlier Democratic-authored opinions than for earlier Republican-authored opinions.

But once we break the cases down into four time periods, we see two different progressions in the point estimates over time (though some are within the reported confidence intervals): first, the magnitude of the partisan point estimates increases; second, the increase over time is greatest for ideologically salient published opinions. In other words, the point estimates suggest that partisanship in general increases over time, and partisanship with respect to the most ideological cases increases the most. And the final period (2007–2017) presents a particularly sharp increase, resulting in partisan differences that are not merely only statistically significant but also large and thus substantively meaningful.

How can we explain that shift over time? We begin by considering whether factors such as a replacement via presidential cohorts or the aging of judges are likely explanations. The data reveal that these factors are unlikely to explain our primary results. We then examine a possibility that is consistent with our findings—different rates of acceleration of ideological change across parties.

1. Partisan Differences Are Not a Function of Presidential or Age Cohorts

One might imagine that part of the story has to do with the increasing politicization of the nomination and confirmation process that started in the Carter and Reagan administrations and carries through to this day.[119] More specifically, one might suspect that the increasing size of partisan differences in recent time periods has something to do with a presidential cohort effect, in which some presidential administrations might have outsized abilities to shift the ideological makeup of the courts. It is widely believed that there are nontrivial ideological differences between presidential cohorts of the same party.[120] If so, those differences might shed light on the partisan differences we find. Perhaps it is not that Democratic and Republican appointees in general are becoming more polarized but that earlier Presidents in our sample appointed moderate Democrats and Republicans and later ones appointed more extreme Democrats and Republicans, with the result that the replacement of the earlier presidential cohorts by the later cohorts produces the polarization we find.

We evaluate that possibility in this section. To summarize, our data do not (somewhat to our surprise) support the claim that the observed increases in partisan differences are due to presidential cohorts, or more generally to the replacement of moderate judges by more extreme judges.

A launching point is the fact that there have been changes in the presidential selection process. For most of the twentieth century, the party of the appointing President was not a particularly strong indicator of ideology. Presidential administrations deferred to a considerable degree to Senators' preferences in choosing circuit nominees, and those Senators often did not prioritize ideological

119. *See infra* notes 123–126 and accompanying text on the roles of Carter and Reagan in changing the process by which judges were chosen.

120. *See infra* notes 123–127 and accompanying text.

commitment to the national party in their choices.[121] Indeed, they were often patronage positions.[122]

The first big move away from senatorial influence came under President Carter, who appointed nominating commissions for each circuit.[123] Those commissions took recommendations from Senators, but the commissions made their own recommendations to the President (to the great annoyance of many Senators).[124] Carter proved to be a way station toward the more complete control that began in the Reagan Administration.[125] Reagan moved to a model of judicial selection that prioritized presidential discretion over senatorial influence, with a small group within the Reagan Administration choosing circuit nominees after engaging in extensive screening that emphasized ideology—"the most systematic judicial philosophical screening of judicial candidates ever seen in the nation's history."[126] Centralized control remained

121. *See* AMY STEIGERWALT, BATTLE OVER THE BENCH: SENATORS, INTEREST GROUPS, AND LOWER COURT CONFIRMATIONS 3–5 (2010). The Eisenhower Republicans on the Fifth Circuit were by many measures more liberal than the Kennedy Democrats, because the former came from the desegregationist party in the South and the latter from the segregationist party in the South. *See also* JACK BASS, UNLIKELY HEROES 84–96 (1981); VICTOR S. NAVASKY, KENNEDY JUSTICE 269 (1971); Kenneth N. Vines, *Federal District Judges and Race Relations Cases in the South*, 26 J. POL. 337, 348 (1964).

122. *See* NANCY SCHERER, SCORING POINTS: POLITICIANS, ACTIVISTS, AND THE LOWER FEDERAL COURT APPOINTMENT PROCESS 13 (2005) (noting that judgeships were "distributed to friends and campaign contributors").

123. A Nixon aide named Tom Charles Huston recommended that Nixon focus on judicial nominations. The memo was forwarded to Nixon's Deputy Attorney General with Nixon's endorsement, but it did not lead to the centralization of the process within the Nixon Administration. *See, e.g.*, Elliot E. Slotnick, *Federal Judicial Selection in the New Millennium*, 36 U.C. DAVIS L. REV. 587, 590 (2003) ("Despite [the Huston] memo, it is a bit too easy to point to the Nixon administration as the historical point in time where the most significant changes took place in the nature of federal judicial selection. The policy implications of judicial selection, which Huston spoke of, were not fully realized until the centralization of the judicial selection process during the Reagan years. More accurately, the modern era of contentious, politicized judicial selection politics can best be traced to the Carter administration.").

124. *See* SHELDON GOLDMAN, PICKING FEDERAL JUDGES: LOWER COURT SELECTION FROM ROOSEVELT THROUGH REAGAN 126–31 (1997).

125. *See* Slotnick, *supra* note 123, at 593 ("Once the genie of openly avowed policy considerations in judicial selection had been let out of the bottle [under Carter], and once the White House's political role in judicial selection increased, it would be difficult to return to the old ways. In the wake of the Carter years, during the two-term presidency of Ronald Reagan, the policy agenda of the president and centralized White House control of judicial selection was a major facet of selection processes.").

126. Sheldon Goldman, *Reagan's Judicial Legacy: Completing the Puzzle and Summing Up*, 72 JUDICATURE 318, 319–20 (1989); SCHERER, *supra* note 122, at 161.

for later administrations, but many researchers suggest that there are substantively meaningful differences in presidential cohorts of the same party—for instance, that the Reagan and George W. Bush judges were more conservative than the George H.W. Bush judges.[127]

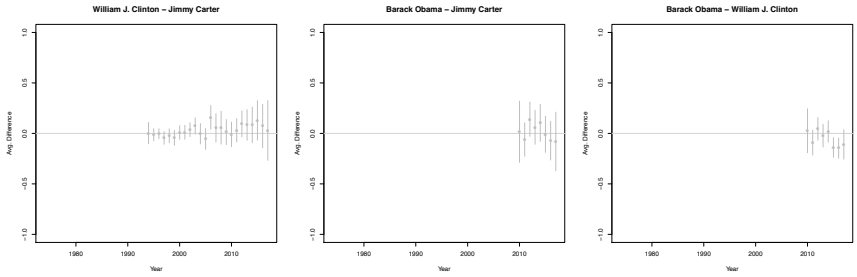
The large size of our dataset gives us leverage to examine the relationship among presidential cohorts. More specifically, our data allow us to examine whether a judge's presidential cohort is associated with that judge's treatments of earlier opinions. We examine this question directly. To shed light on the proposition that increasing political polarization is linked to selection of judges, we look at substantive treatments within a cohort on a year-by-year basis to see whether any cohort-specific differences become larger over time.[128]

We begin by examining the intraparty differences between presidential cohorts with respect to the propensity of judges from each cohort to follow opinions written by Democratic authors and Republican authors. We examine this by looking at differences in the average number of positive treatments coming from judges within the various intraparty cohorts after adjusting for circuits. *Figure 9* displays these differences over time for the three pairings of Democratic presidential cohorts.

127. *See, e.g.*, SUNSTEIN ET AL., *supra* note 7, at 113 ("By common lore . . . President Reagan was determined to 'stock' the federal bench with conservative judges, whereas President George H. W. Bush was significantly more moderate and President George W. Bush behaved more like President Reagan.").

128. It is worth noting that the analysis of cohorts faces a fundamental identification problem. Namely, if there are age effects (effects specific to judges of a certain age regardless of calendar year or presidential cohort), period effects (effects that are present for all judges in a particular calendar year regardless of cohort and age), and cohort effects (effects specific to a cohort of judges regardless of the judge's age or the calendar year), it is impossible to separate these effects. *See* Willard L. Rodgers, *Estimable Functions of Age, Period, and Cohort Effects*, 47 AM. SOCIO. REV. 774 (1982). For instance, if cohorts are defined by a judge's birth year, then calendar year – age = cohort, which results in perfect collinearity among the right-hand-side variables of a regression. Thankfully, we do not face this severe problem. Because we are defining cohorts in terms of each judge's appointing President, and because Presidents nominate judges of multiple ages within and across multiple years, our age, calendar year, and presidential cohort variables are not perfectly collinear. Nonetheless, these variables are correlated, which creates challenges for inference.

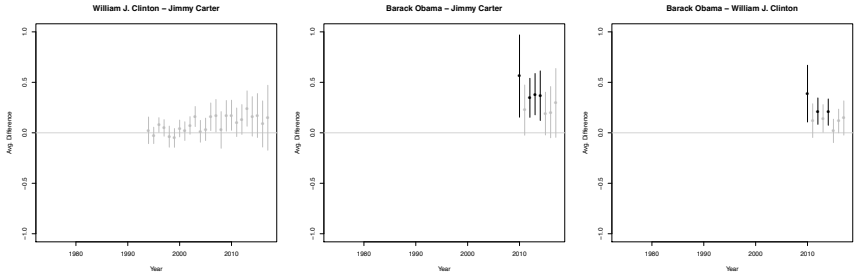
Figure 9



Differences in the average number of positive treatments to opinions authored by Democratic appointees by Democratic presidential cohorts. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are in black. Statistically insignificant differences are in gray.

We see that there are effectively no differences between Democratic presidential cohorts in terms of their propensity to follow previous opinions authored by Democratic appointees.

Figure 10



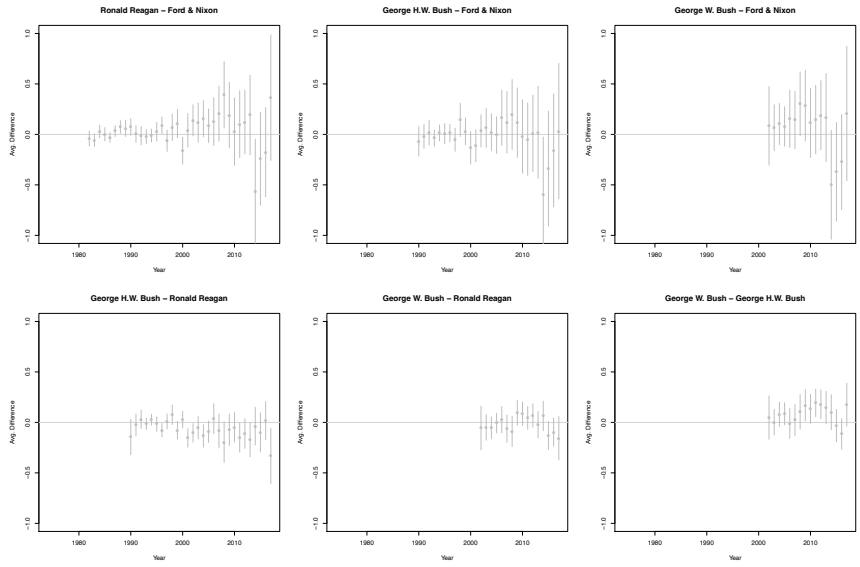
Differences in the average number of positive treatments to opinions authored by Republican appointees by Democratic presidential cohorts. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are in black. Statistically insignificant differences are in gray.

Figure 10 plots similar results comparing the extent to which different Democratic presidential cohorts follow opinions written by Republican appointees. Again, most of the year-by-year differences are not significantly different from zero. The few differences that are significant occur in the 2010s where there is some evidence that the Obama cohort is more likely to follow Republican-authored opinions than the Carter cohort and, to a lesser extent, the Clinton cohort. Note that this pattern of more

recent appointees being more likely than earlier appointees to positively treat opinions authored by judges of the opposing party is exactly the opposite of what we would expect if the increased politicization of the nomination and confirmation process is driving increases in partisan differences.

We repeat the same analysis for the Republican presidential cohorts. Again, we look at differences in the average number of positive treatments coming from judges within the various intraparty cohorts after adjusting for circuits.

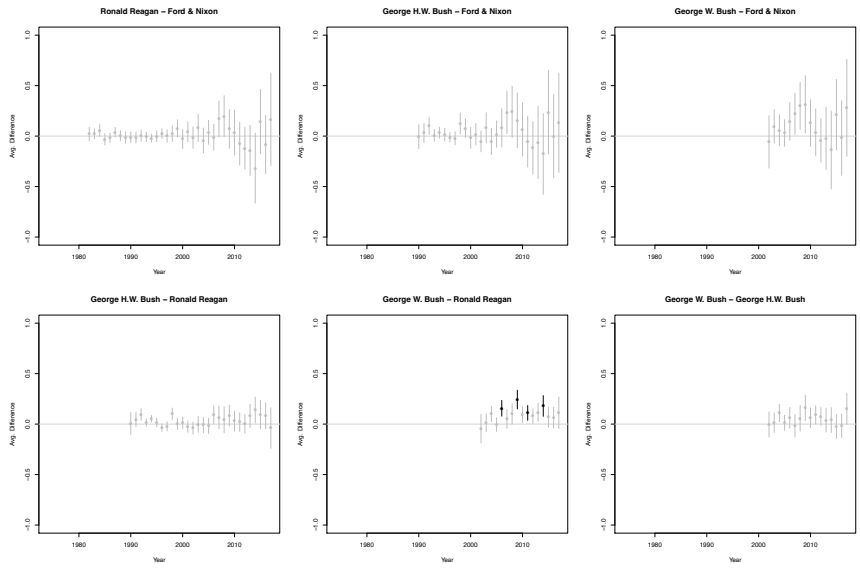
Figure 11



Differences in the average number of positive treatments to opinions authored by Republican appointees by Republican presidential cohorts. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are in black. Statistically insignificant differences are in gray.

Figure 11 displays the resulting differences over time in how six pairings of Republican presidential cohorts positively treat earlier Republican-authored opinions. As we saw with the Democratic presidential cohorts above, there is no statistically significant difference between any of the Republican presidential cohorts.

Figure 12



Differences in the average number of positive treatments to opinions authored by Democratic appointees by Republican presidential cohorts. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are in black. Statistically insignificant differences are in gray.

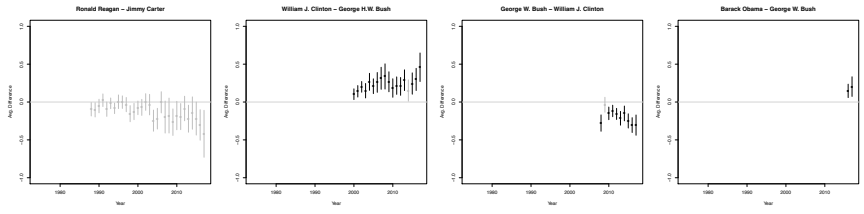
Figure 12 reports similar results for the differences in how these six pairings of Republican presidential cohorts positively treat opinions authored by Democrats. The vast majority of these year-specific differences are not statistically different from zero.

Taken as a whole, the evidence in *Figures 9–12* paints a picture in which there are few, if any, differences among intraparty cohorts in how earlier opinions are substantively treated. *Figures 9–12*, together with *Figures 1* and *2*, indicate that there are large, substantively meaningful differences between Democratic and Republican appointees but very small differences among Republican appointees and among Democratic appointees. These results are consistent with a provocative claim: scholars may attribute too much significance to intraparty presidential cohorts. Partisanship matters, but within a given party the identity (and ideology) of the appointing President does not. This goes against most of the conventional wisdom on the importance of presidential cohorts within the same party.

An issue with the analyses in *Figures 9–12* is that judges are entering and leaving the court at different times. Accordingly, the over-time comparisons involve different judges. To address that potential issue, we next look at cross-party comparisons of temporally adjacent presidential cohorts that are also restricted to only those judges who served from the last year of the second presidency in each pairing to 2017. By holding the set of judges fixed while also adjusting for circuit we hope to eliminate any changes in outcomes due to changes in the pool of judges being compared.

Since we are now looking at cross-party differences in substantive treatments, we expect to see statistically significant differences. However, the real question involves the trajectory of those cross-party differences in behavior. If the differences are constant over time, with the differences between some pairs much larger than others, this would be consistent with presidential cohort effects. On the other hand, if the differences we see tend to be trending over time and largest in the most recent periods, this would be more consistent with increasing partisan polarization—driven not by presidential-cohort fueled replacement on the bench (since we are looking at the same judges throughout the period covered in the figures) but rather by factors related to increasing political polarization in the country.

Figure 13



Differences in the average number of positive treatments to opinions authored by Democratic appointees for temporally adjacent cross-party presidential cohorts. The judges included are only those from the cohort pair in question who served for the entirety of the period from the last year of the second presidency in each pairing to 2017. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are in black. Statistically insignificant differences are in gray.

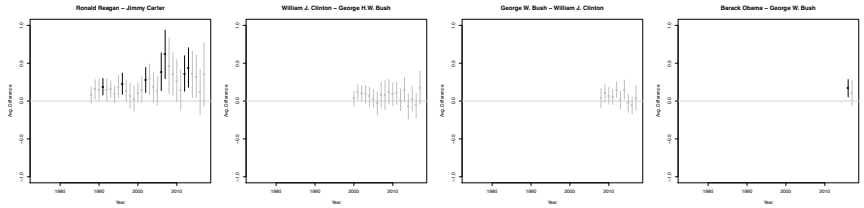
Figure 13 plots the cross-party differences in the average number of positive treatments to Democratic-authored opinions for these four temporally adjacent cross-party cohorts. Note that we

see significant differences for three pairs of adjacent cohorts. In each case, the differences are in the expected partisan direction. Note that these differences appear to be trending away from zero (at least for the pairs with more than two years of data). Further, the Clinton-H.W. Bush plot is essentially the mirror image of the W. Bush-Clinton plot for the years 2008 to 2017 when there are data for both pairs. All this is not surprising given our finding of similarities among presidential cohorts within parties: given that presidential cohorts within parties are similar, we would expect to see roughly similar trends when we compare the appointees from any two presidential cohorts from different parties. Again, this is more consistent with a partisan polarization story than a presidential cohort story.

The one exception is surprising: there are no statistically significant differences between the Reagan cohort and the Carter cohort for any year from 1988 to 2017. Given the primacy that the Reagan administration's approach to judicial appointments has in the conventional wisdom regarding the politicization of the judiciary, we might have expected large differences between the Carter and Reagan cohorts. We find none. This is all the more interesting given that the Carter administration is viewed by some scholars as the actual starting point for the politicization of judicial appointments.[129] That said, while there are no statistically significant differences, the point estimates do trend toward a partisan difference in the expected direction—and, importantly, this trend begins only in recent years. This is suggestive of partisan polarization rather than a presidential cohort effect; however, it is important to emphasize that this trend is not statistically significant at conventional levels.

129. See *supra* notes 123–125 and accompanying text.

Figure 14



Differences in the average number of positive treatments to opinions authored by Republican appointees for temporally adjacent cross-party presidential cohorts. The judges included are only those from the cohort pair in question who served for the entirety of the period from the last year of the second presidency in each pairing to 2017. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are in black. Statistically insignificant differences are in gray.

We can also look at how these adjacent cohort pairs compare in their propensity to positively treat opinions written by Republicans. This information is displayed in *Figure 14*. In some sense, these results are even easier to interpret, as nearly all the within-pair differences are not statistically different from zero. Once again, we see very little difference between the Carter judges and the Reagan judges (only one difference is significantly different from zero prior to 2007). At that point the differences do trend up in a partisan direction and two become significant. Again, this is more consistent with partisan polarization in the 2000s than with presidential cohort effects.

To save space, we present related aggregated analyses in Appendix A4. These results compare particular Democratic presidential cohorts to all Republicans and particular Republican presidential cohorts to all Democrats. These results tell a similar story to those above in *Figures 13* and *14*.

An alternative explanation is that what we are assuming is a partisan polarization effect could be an age effect. Perhaps judges become more ideologically extreme (or at least less likely to positively treat opinions by judges of the opposing party and more likely to positively treat opinions by co-partisans) as they become older. To investigate this, we examine the cross-party difference in the average number of positive treatments to Democratic-authored opinions within judges born in the same year over time as well as the cross-party difference in the average number of positive treatments to Republican-authored opinions within judges born in

the same year over time. These analyses also adjust for circuit. The resulting plots are displayed in Appendix A3 for reasons of space. The key takeaway from this analysis is that within judges with the same birth year there are very few year-specific partisan differences in the propensity to positively treat either Democratic or Republican opinions. The within-birth-year partisan differences that do emerge are very largely in the years after 2000. These differences are not found predominantly among the earlier birth-year cohorts. Indeed, the one birth-year group that exhibits a consistently significant partisan difference in the average number of positive treatments to Democratic opinions after about 2004 is the group of judges born in 1948, which is about halfway between the earliest birth-year group (1926) and the most recent (1967). As with presidential cohorts, partisan polarization in the 2000s is a more compelling explanation than an account based on aging.

2. The Best Explanation: An Accelerating Move Right Among Republicans but a Steady Move Left Among Democrats

Given that explanations focusing on presidential cohorts or the aging of the judges do not hold water, how might we explain the growing magnitude of partisan differences in our data? Some form of fairly recent ideological divergence between Democratic and Republican appointees seems like the best answer. This implicates another finding that surprised us: the partisan differences in treatment of Democratic-authored opinions and of Republican-authored opinions are both statistically significant, but they are larger for Democratic-authored opinions. At first blush, this asymmetric pattern of behavior does not seem consistent with simple explanations rooted in ideology. What we might call the simple ideological account is the view that judicial behavior depends only on the relative ideological locations of current judges. According to this simple ideological account, sitting Democratic and Republican appointees should behave as mirror images of each other. We do not see that in the data, so this simple ideological account cannot be correct.

What explains this pattern of larger partisan differences for the treatment of Democratic-authored opinions than Republican-authored opinions? In general, focusing on attributes and characteristics of the later judges does little to explain this pattern. If some attribute of later judges distinguished later Republican

appointees from later Democratic appointees, presumably such a distinction would apply equally to their treatment of earlier Democratic-authored opinions and Republican-authored opinions. For example, the possible increasing partisanship of the later judges would not explain the finding: if, say, the Republicans and/or Democrats on the later panels were more partisan, then we would expect that divergence to show up in the treatment of earlier Democratic-authored opinions and earlier Republican-authored opinions.

We believe that the key variable must be some difference in the earlier Democratic-authored opinions compared to the earlier Republican-authored opinions—or, more to the point, in the later judges' perception of those earlier opinions. What aspect of these earlier opinions might be driving the result we observe? While we cannot be sure what the relevant aspect is, we do have strong reasons to exclude some things from consideration.

One might be tempted to think that the relevant difference between these earlier Democratic- and Republican-authored opinions has something to do with the types of cases heard by panels with different partisan compositions. But the judges and cases are randomly assigned to panels within a circuit and year, so there should not be any such differences within a particular circuit and year.

Nonetheless, one might wonder about the stock of earlier opinions that can be followed and whether partisan imbalances in the circuit in question might create more opportunities for selective treatment of the Democratic-authored opinions than the Republican-authored opinions. This is unlikely. Note that the median time difference between the treating and treated opinion is 5.1 years and the mean time difference is 7.5 years. In other words, most positive treatments are to opinions that were issued relatively recently. Further, recall that our estimation approach weights the data so that treating opinions that are equally likely to be authored by each side of the comparison (e.g., as likely to be authored by a Democrat as a Republican) get the highest weight.[130] The weights go to zero as the likelihood of each type of author becomes increasingly unequal. Finally, note that the partisan composition of most circuits changes fairly slowly over time. Taken together, these three facts

130. *See* Section IV.C.

mean that the treating cases that receive the most weight in our analysis have a stock of cases to follow that is composed of a roughly similar mix of opinions written by Democratic and Republican appointees.

Other possible differences in earlier opinions do not have much explanatory force. For instance, if we imagine that earlier Republican-authored opinions were more attractive for some nonpartisan reason (e.g., they had better reasoning), that would not explain why there is a divergence in how later judges treat Democratic-authored versus Republican-authored opinions. If later Democratic and Republican appointees recognized that better reasoning and deemed it important, then presumably they would similarly treat (well) the earlier Republican-authored opinions and similarly treat (poorly) the earlier Democratic-authored opinions.

How, then, can we explain the differences in the substantive treatment of Democratic-authored opinions versus Republican-authored opinions? We surmise that Republican-authored opinions were perceived as more equally acceptable to later Republican and Democratic appointees than Democratic-authored opinions were. The most obvious possible basis for a difference in acceptability is a difference in *perceived ideological distance of the earlier opinions from the preferences of later judges.* If later Republicans became dramatically more conservative but later Democrats did not become dramatically more liberal, then we might expect to see the pattern in our data.[131] Later Democrats would see significant differences between earlier Democratic-authored and Republican-authored opinions and would find earlier Democratic-authored opinions much more acceptable. Later Republicans would find earlier Republican-authored opinions slightly more acceptable than earlier Democratic-authored opinions but would find neither set particularly attractive because (by hypothesis) the later Republicans would have become so much more conservative.

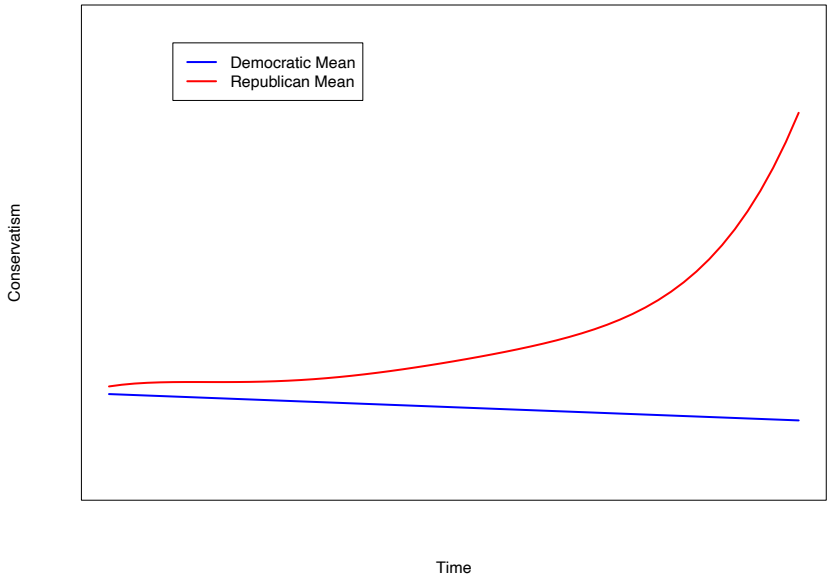
The result would be that later Democrats and Republicans would respond differently to earlier Democratic-authored opinions (because later Democrats would find earlier Democratic-authored opinions quite attractive, and later Republicans would find earlier Democratic-authored opinions unattractive); but later Democrats

131. Note that such party-specific differences in ideological movement are consistent with accounts of ideological polarization in Congress. *See, e.g.*, DeSilver, *supra* note 12; Lewis, *supra* note 12.

and Republicans would respond more similarly to earlier Republican-authored opinions (because later Democrats would find them unattractive, and later Republicans would find them only marginally more attractive than later Democrats do, because Republicans would be so much more conservative than their Republican predecessors that these earlier Republican-authored opinions would not be significantly more attractive). What separates this nuanced ideological account from the simple ideological account discussed above is that the ideology of earlier judges (and by extension their opinions) enters the explanation, whereas the simple account looks at only the ideology of current judges deciding whether to follow earlier opinions.

To investigate our surmise that partisan differences in ideological distance from earlier Democratic-authored and Republican-authored opinions might explain our findings with respect to those opinions, we created a simulation model. The model assumes that when deciding which earlier opinions should be followed, judges look back to earlier opinions within a temporal window with the probability of a positive treatment being a decreasing function of the distance between the treating judge's ideology and the earlier opinion author's ideology. Simply stated, the model assumes that more ideologically proximate authors are more likely to be followed. The model assumes that Democratic appointees, on average, have different ideologies from Republican appointees. It also assumes that these ideological differences change over time. Specifically, it assumes the average ideology of a Republican appointee is distinct from the average ideology of a Democratic appointee and that Republican appointees are becoming exponentially more conservative over time while Democratic appointees are trending in the opposite direction in a linear fashion. *Figure 15* displays this hypothetical ideological divergence.

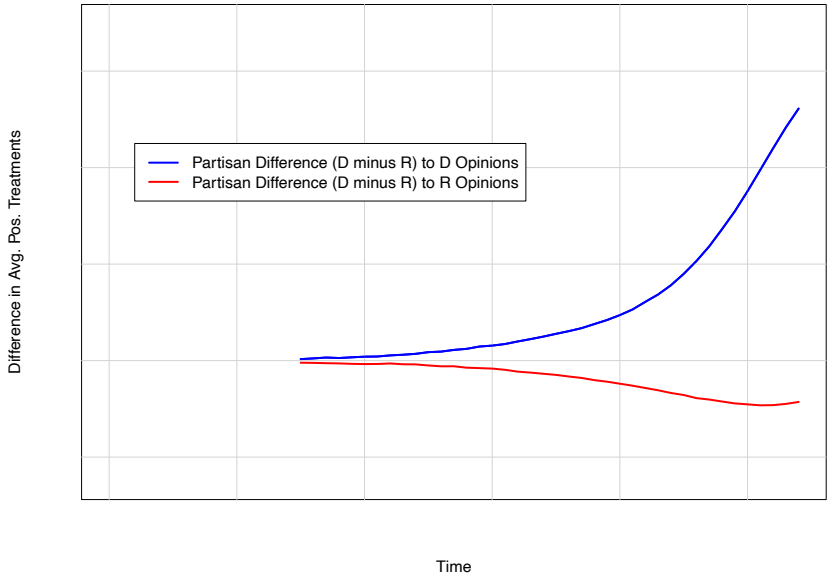
Figure 15



Hypothetical change in the average ideological location of Democratic and Republican appointees over time. This is the divergence that is posited in the simple simulation study discussed in the text. The key point that drives the results is the accelerating change among Republican appointees compared to the steady change of Democratic appointees.

Putting these assumptions together and letting our simulated judges make simulated decisions to follow opinions produces the partisan differences in positive treatments to Democratic-authored opinions and Republican-authored opinions depicted in *Figure 16*. Note that the partisan difference in the propensity to follow Democratic-authored opinions increases rapidly over time. The partisan difference in the propensity to follow Republican-authored opinions grows more slowly than the partisan difference to follow Democratic-authored opinions and eventually stabilizes. This pattern largely mirrors the average controlled differences by author partisanship over time in *Figure 2*.

Figure 16



Partisan differences in positive treatments of Democratic-authored opinions and Republican-authored opinions produced by the simulation model discussed in the text.

What is the intuition behind these results? The crucial step is the assumed accelerating movement in a conservative direction among Republican appointees compared to a steady movement in a liberal direction among Democratic appointees. The result is that Republican-authored opinions decided considerably earlier are roughly halfway between the positions of current Democratic appointees and current Republican appointees. That is, such temporally distant Republican opinions are roughly equally attractive to current Democratic appointees and Republican appointees—although more recent Republican opinions will still be substantially more attractive to Republicans than to Democrats. This will attenuate the partisan differences in the propensity to follow Republican opinions. By contrast, because Democratic appointees are not trending to the left at an increasing rate, there will be much less attenuation of partisan differences in the propensity to follow Democratic-authored opinions. If Republican appointees are becoming exponentially more conservative over time and Democrats' move to the left is only linear, then Republican appointees will be only mildly more positive about older

Republican as opposed to older Democratic opinions, whereas Democrats will have a stronger preference for Democratic over Republican opinions. If so, then we would expect exactly the results that the model suggests in *Figure 16* and that we find in this Article.

None of this is to say that Republican appointees have become, in some objective sense, more ideologically extreme than Democratic appointees. The starting location prior to the ideological divergence is arbitrary. It is consistent with a world where both Democratic and Republican appointees start off on the extreme left in some objective sense and the Republican appointees are tracking back to moderate positions at an accelerating pace while Democratic appointees move to more extreme left positions at a steady rate.

So our data suggest that there has been a rapidly accelerating divergence in the relative conservatism of appeals court judges across parties and that this accelerating divergence is driven mainly by an accelerating move to the right by Republican appointees. While we did not set out to provide direct measures of judicial ideology for the federal courts of appeals, our analysis does provide indirect evidence of such an accelerating divergence in judicial ideology, driven primarily by an accelerating rightward move by Republican appointees. Absent a more compelling alternative explanation, which we have not been able to articulate, we believe this accelerating divergence is the best explanation for the differences we find in the treatments of Democratic-authored opinions versus Republican-authored opinions.[132]

132. The recent rise in partisan differences raises a possibility suggested by our colleague Maggie Lemos: Could it be that there was a significant broadening in access to Lexis and Westlaw immediately preceding the rise, such that judges would have newly gained access to these vast online databases and thus were able, for the first time, to comprehensively look for earlier opinions written by ideologically compatible judges? The answer is no, for two reasons. First, Lexis and Westlaw sent us chronologies of their availability, and both were widely available to judges by the early 1980s, with full coverage of circuit court opinions. Second, this could not explain the recent divergence in Republican and Democratic appointees' treatments of earlier Democratic-authored opinions but not earlier Republican-authored opinions. Insofar as the availability of Lexis and Westlaw allowed Democratic appointees to do a better job of finding Democratic-authored opinions that they could follow, or allowed Republican appointees to do a better job of finding ways not to follow Democratic-authored opinions, we would expect to see a similar pattern with respect to Republican-authored opinions—and we do not.

A. Racial Differences

Our investigation of the relationship between a judge's race and positive treatments of opinions by Black and Hispanic authors reveals patterns that are fairly similar to our findings on partisanship, particularly for opinions by Black authors. And our results are not consistent with the solidarity effect that we posited in section III.C.

Starting with opinions by Black authors, *Figure 3* indicates that across the entirety of the time period Black judges are more likely than White judges to follow opinions written by Black co-partisans. The differences are statistically significant and comparable in size to the partisan differences discussed above. Breaking the data into four time periods in *Figure 4*, we see progressions over time that resemble those for partisan differences. First, the magnitude of racial differences increases. Second, the increase over time is greatest for ideologically salient and race-salient published opinions. And the final period presents a particularly sharp increase. In that final period, the differences are statistically significant and substantively meaningful, and they are particularly large for opinions authored by Black Democratic judges.

The White-Hispanic differences we see are more muted than the White-Black differences. When we pool the data over the full time span of the study, the differences are statistically significant (except for race-salient opinions by Democratic authors) but substantively small. The differences are smaller, and thus less meaningful, for opinions by Hispanic Democratic authors. It is particularly interesting to note that the difference between how White and Hispanic Democratic opinion authors treat opinions written by Hispanic Democrats in race-salient cases is not significantly different from 0—suggesting that racial/ethnic identity is not driving the White-Hispanic differences (at least within Democrats).

Turning to the four time periods, for Hispanic Republicans the clearest pattern is that the differences increase over time. As for different types of cases, there are relatively modest differences in the point estimates (and all are well within the confidence intervals). For opinions by Hispanic Democrats, the differences broken into different time periods are quite small and generally are not statistically significant. There is a modest upward trend in the point estimates over time for ideologically salient and race-salient

cases, but given the lack of statistical significance we do not draw any conclusions from it.

Relatedly, as we noted above, the relatively small number of Black (and the even smaller number of Hispanic) judges means that our results are based on fewer cases from an idiosyncratic selection of circuits with much more weight given to recent cases. Because 71.2% of *Shepard's* treatments occur within circuits, comparisons depend on sufficient numbers of judges with the relevant attribute in a given circuit, and there are fewer circuit-year combinations that have Black or Hispanic judges than have female judges. As we also noted above, our decision to focus on average controlled differences with overlap weights (and conditioning on circuit, year, and circuit-year combinations to avoid spurious results) further serves to downweight data from many circuits and years. To be clear, this downweighting of data from circuit-years where good apples-to-apples comparisons are not available is a strength of our research design, but it does mean that our results need to be interpreted with care.

The fact that we see intraparty racial differences is notable because it indicates behavioral differences among judges of the same party along racial lines. The Black-White racial differences (particularly for Black Democratic authors, who are more numerous than Black Republican authors) are particularly striking: Black Democratic appointees significantly diverge from their White Democratic counterparts in their treatment of cases written by Black Democratic authors, and to a lesser extent Black Republican appointees differ from White Republican appointees in their treatment of opinions by Black Republican authors, even though in both cases we are comparing co-partisans' treatment of their fellow partisans' opinions. A possible explanation for the smaller differences for Hispanic Democratic judges relative to White Democratic judges is that Hispanic Democratic judges may not be appreciably more liberal than White Democratic judges.[133] Given the small number of Hispanic Republican judges, the presence of statistically significant results for them is surprising. But the small number of such judges also means that our results reflect the decisions of a relatively small number of judges, giving us less

133. *See* Sen, *supra* note 15, at 394 tbl.4 (finding, based on preconfirmation campaign contributions, that Hispanic Democratic judges are ideologically closer to White male Democratic judges than are female or Black Democratic judges).

confidence with respect to Hispanic Republican judges than for Black judges that we have uncovered an important pattern.

So what to make of these intraparty racial differences, particularly for opinions by Black authors? The data are not consistent with the solidarity effect that we posited in section III.C. As we noted there, a solidarity effect would presumably either be constant or decrease over time. Either Black judges would prefer to follow opinions issued by another Black judge, or trailblazing Black judges might prefer to support other Black judges when they were very few in number and struggling to be accepted. The first possibility would lead to a constant effect and the second would likely lead to a diminishing effect in the most recent time period, or perhaps also a constant effect (insofar as current Black judges see themselves as trailblazers).[134] These theories are not consistent with the results we find, where the observed differences rise dramatically in the most recent time period. Rather than evidence for solidarity effects, we instead see the trend that we also saw with respect to partisan differences.

By contrast, the general ideological and more policy-specific hypotheses we discussed in section III.B seem to fit well with our results. We think that the data better fit an explanation that places more of the weight on broad ideology than on expertise and lived experiences, for two main reasons. First, recall that race-salient cases are a subset of ideologically salient cases, because the categories that researchers have identified as race-salient have also been identified as ideologically salient.[135] The key difference is that the subset of race-salient cases, in addition to reflecting ideology generally, may reflect differences in lived experiences. The subset of ideologically salient cases that does not include race-salient cases is a purer measure of ideology (i.e., without the additional element of expertise and lived experiences). Because this subset of ideologically salient cases excludes the cases for which lived experiences have been posited as relevant, the large difference for this subset indicates that ideology is playing a large role. The fact that the differences in race-salient cases are statistically indistinguishable from the differences in the subset of ideologically

134. *See supra* notes 93–94 and accompanying text.

135. *See supra* note 82. Gender-salient cases are also a subset of ideologically salient cases. *Id.*

salient cases that exclude race-salient cases suggests that lived experience is, at most, playing a modest incremental role beyond that played by general ideology.

Second, the rise in the differences between Black and White judges follows the same basic timing pattern as the differences based on partisanship: for both race and partisanship, the differences rise over time, particularly in the most recent time period. As we noted above, what the category of ideologically salient cases is designed to measure is similar to what partisanship measures.[136] In both cases, we are measuring ideology broadly understood. The striking similarity of the rise in partisan differences and differences in ideologically salient cases suggests that there is a commonality between them in the form of broad ideology.

It is of course possible that the rise in partisan differences happened to coincide with a rise in intraparty racial differences in relying on expertise and lived experiences, but we put less weight on such a possibility. We can think of three reasons why the two might coincide. First, it might be that race-salient cases are connected to partisanship such that the rise in both is not just a coincidence. Insofar as the race-salient category measures differences in expertise and lived experiences, it is not clear what that connection would be. The theory behind the emphasis on expertise and experiences, after all, is that it is different from ideology. This relates to a second possibility: it might be that the category of race-salient cases in fact measures ideology akin to partisanship, rather than measuring expertise and lived experiences, such that both reflect a rise in ideological differences. But, if so, then this is just an ideological story after all: the category of race-salient cases would in fact be measuring ideology broadly understood. Third, it might be that the rise in partisan differences and in intraparty differences for race-salient cases is just a coincidence. We cannot rule out that possibility, of course. But given the connection between what partisanship measures and broad ideology (as measured by ideologically salient cases), we think a correlation between partisanship and broad ideology is more likely than the happenstance of two unrelated phenomena showing the same pattern over the same period of time.

136. *See supra* text accompanying notes 84–86.

The fact that the recent rise in intraparty racial differences accords with the rise in partisan behavior does not mean that we have proved that ideology is the explanation for racial differences—or partisan differences, for that matter. For both race and partisanship, there may be a non-ideological explanation for our results. But the most obvious non-ideological explanation is a solidarity effect, and our data are not consistent with such an effect.

We think the best explanation is that this increase in racial differences is indeed akin to the increase in partisan differences, reflecting the same trend—greater ideological polarization over time. Insofar as race may reflect components of ideology that partisanship does not reflect, increasing racial differences likely reflect aspects of ideological polarization that partisanship does not capture. Simply stated, it appears that judges' substantive treatments of their colleagues' opinions are reflecting greater polarization that is manifested in both partisan and intraparty racial differences.

The implications of these results relate to broader issues of race and ideology. As we noted in section II.B.ii, previous researchers have attempted to identify differences in judicial behavior between Black and White judges by examining their votes in a variety of case types, and they have found differences in relatively few categories of race-salient cases.[137] The null results in most categories of cases have been interpreted as indicating that there are no significant differences in the behavior of Black and White judges outside of this small number of race-salient categories.[138] But, as we have noted in this Article, votes are only one form of judicial behavior. Our dataset allows us to carefully examine racial differences in substantive treatments, and our findings indicate that there are substantively meaningful racial differences in behavior across a wide range of cases. By looking at substantive treatments, we have found differences that others have missed.

137. *See supra* notes 62–66 and accompanying text.

138. *See, e.g.*, Burbank & Farhang, *Pleading Decisions*, *supra* note 51, at 2246 (canvassing the null results outside of sex discrimination in employment and concluding: "In sum, the employment discrimination studies revealing gender differences in Court of Appeals decision-making are islands in a sea of null results.").

Our results suggest that Black judges ideologically differ from their White co-partisans.[139] And those differences are consistent with the idea that Black judges are, on average, to the left of White judges. Party is a measure of ideology. It appears that race can refine that measure.

The results for differences between Hispanic and White judges are less meaningful, for two reasons. First, the results for Democratic appointees are small and not substantively meaningful (and often statistically insignificant). Over the full 1974–2017 span of our study, the White-Hispanic average controlled differences are always substantively small. Further, the White-Hispanic average controlled difference within the subset of race-salient cases is not different from 0. And when we break down the results into our four time periods, most of the results for Hispanic Democrats lack statistical significance. All of this points to limited influence of Hispanic identity on Democratic appointees' following of earlier opinions.

Second, the White-Hispanic differences we report are based on limited data. Our decision to focus on average controlled differences with overlap weights that adjust for circuit, year, and circuit-year means that circuit-year combinations that have no Hispanic judges do not enter into our analyses. Even circuit-years with some Hispanic judges may get very little weight if there is, say, only one Hispanic judge in that circuit-year. In the time span of our study, nine circuits never had a Hispanic Republican judge, and only two had more than one. Practically, this means that the White-Hispanic comparisons we report are best thought of as very localized comparisons of White and Hispanic judges in those (relatively few) circuit-years with multiple Hispanic judges and White judges. Indeed, our results with respect to White and Hispanic Republican appointees derive, to a very large extent, from three circuits—the First, Fifth and Ninth. Our results reflect the universe of meaningfully comparable White and Hispanic Republican appointees, but that universe is a very limited one.

Although we place little weight on these results, the pattern over time for Hispanic Republicans is fairly similar to those for Black judges (and, as we shall see, for female Republican judges):

139. And that Hispanic Republicans ideologically differ from White Republicans, and are, on average, to the left of White Republicans.

differences become larger over time, with the differences in the 2007–2017 time period being quite large. The confidence intervals for all the types of cases in the last time period have substantial overlap, suggesting that the differences we find do not depend appreciably on the subset of cases examined. But these findings can be taken as suggestive of substantively meaningful differences between White and Hispanic Republican judges. And, as with differences between Black and White judges, these differences are consistent with the proposition that Hispanic Republicans are, on average, less conservative than White Republicans.

C. Gender Differences

Turning to gender differences, we see in *Figure 7* that, pooling the data across the entirety of the 1974–2017 time period, female judges are more likely than male judges to follow opinions authored by co-partisan women. The differences are statistically significant, but they are not large and not substantively meaningful. Breaking the data into four time periods in *Figure 10*, the results remain substantively small, but we see different patterns for Republican and Democratic judges. For Republican judges, there are increasing gender differences over time, particularly between the second and third time periods, and particularly for ideologically salient cases (and less so for gender-salient cases). For Democratic judges, the results are less clear: many of the results are not statistically significant, and most of the increases over time are well within the confidence intervals and thus not substantively meaningful. Interesting in this regard is the absence of statistical significance for the ideologically salient and gender-salient cases for three of the time periods. The analogous measures for A) gender differences for Republicans, B) Black-White differences, and C) Hispanic-White differences for Republicans were generally statistically significant (and changed the most over time). But for Democratic gender differences, the main statistical significance occurs in all cases and all published cases.

So, what should we make of these gender results? If we focus on Republican judges, the pattern is fairly similar to the patterns for partisan differences, Black-White differences, and Hispanic-White differences for Republicans (though the gender differences are considerably smaller): the differences increase over time and are greatest for the most ideologically salient cases.

This increase over time for Republican judges is not consistent with the solidarity effects we outlined in section III.C. As we noted in the previous section, a solidarity effect would presumably either be constant or decrease over time. The observed increase in gender differences over time, like that for Black-White differences, and Hispanic-White differences for Republicans, suggests that something other than solidarity effects is the explanation.

Also notable is the fact that the gender difference for Republican judges is not higher for gender-salient cases than for ideologically salient cases or for ideologically salient cases excluding gender-salient cases. Indeed, the latter two categories are statistically significant for all three time periods with sufficient data, and the gender-salient cases are significant only for the 1996–2006 time period. So the case for placing most of the weight on broad ideology rather than expertise and lived experiences is even stronger for the gender differences for Republicans than for Black-White differences. All the same points we made in section VI.B apply, and here the lack of statistical significance for gender-salience in two of the time periods further weakens the case for the gender-salient cases doing much work. All this suggests that the gender differences are better explained by ideological differences, as opposed to the lived experience differences that the gender-salient cases are designed to measure.

The discussion in this section so far has focused on gender differences among Republicans. But when we turn to gender differences among Democrats, the picture becomes much murkier. As we noted above, if we pool the data among all four time periods, there are statistically significant, but small, gender differences. When we break them up into the four time periods, there is a fairly clear pattern of increases in gender differences for all cases. And there are statistically significant differences in all three time periods with a meaningful number of female judges for all and all published cases, and in two of the time periods for ideologically salient cases minus gender-salient cases, but in only the 1996–2006 period for ideologically salient and gender-salient cases (and we do not see a consistent rise in gender differences for ideological or gender-salient cases). The point estimate for gender-salient cases in the 1996–2006 period is a bit higher than for ideologically salient and ideologically salient minus gender-salient cases, but all are well within the confidence intervals. Given the weakness of these

results, it would be a stretch to attempt to draw any meaningful conclusions about the role of broad ideology versus expertise and lived experiences, or the role of any solidarity effect. There simply are not enough data to draw strong conclusions. So if we consider the gender differences for Democrats in isolation, the results are not very substantively meaningful: there are statistically significant differences over the entire time period, but when we break the results into time periods the results are more often statistically insignificant or small and not substantively meaningful.

The murkiness of the gender results for Democrats has implications for the gender results overall. At a minimum, combining the Democratic and Republican gender results means that we cannot identify meaningful consistent patterns about the role of gender (whereas we can identify such patterns for partisanship and for Black-White differences). There is no requirement that we consider the Democratic and Republican gender results together, but we did not hypothesize different gender differences for Democrats versus Republicans.

The most we can say is that the gender results for Democratic judges are not particularly meaningful, but the results for Republicans are suggestive of ideological differences along gender lines akin to those that we saw for racial differences: gender (for Republicans) seems to reveal a growing ideological separation over time. That is, for Republicans gender may illuminate ideological differences that party membership does not reveal. And those gender differences are consistent with the proposition that female Republican judges are, on average, less conservative than male Republican judges.

An ideological explanation for the gender differences among Republicans suggests a possible refinement to the partisan explanation at the end of section VI.A. Recall that our explanation for greater partisan differences for opinions by Democratic judges than for Republican judges involved great Republican rightward acceleration compared to steadily leftward movement among Democrats. The gender differences among Republicans raise the possibility that all Republicans did not move equally quickly in a conservative direction, and that in fact male Republicans accelerated rightward more rapidly than did female Republicans. The Black-White Republican differences suggest that White Republicans accelerated in a conservative direction relative to Black

Republicans. And although the number of observations is smaller and thus the data are less informative, the change over time for Hispanic-White Republican differences is consistent with White Republicans moving rightward more rapidly than Hispanic Republicans. The data are only at best suggestive. But they do provide some mild support for the proposition that an accelerating rightward movement among White male Republican judges had the greatest contribution to the recently rising differences we find in this Article.

CONCLUSION

If we had written this Article twenty years ago, we would have been able to tell a heartening story about ideology—that it was not a significant factor in judges' decisions to follow earlier opinions. But today the story is different: In recent years, polarization has risen along party lines, within both parties for Black versus White judges, and within Republicans for Hispanic versus White and female versus male judges. And these differences are greatest in the most ideologically salient cases. The best explanation is ideological polarization between and within political parties.

And there is an interesting wrinkle: The rising partisan polarization we find in judges' treatment of earlier opinions is greater when a Democrat wrote the earlier opinion. That pattern is consistent with Republican appointees rapidly becoming more conservative and Democrats appointees more slowly becoming more liberal.

More broadly, our data indicate that judges were not particularly ideological in their substantive treatments of earlier opinions in earlier decades, but in the twenty-first century have become increasingly so, with attendant impacts on the shape of legal doctrines. Judges, it seems, are subject to the polarization that affects the rest of us.

APPENDIX

A1. CASE COUNTS ON IDEOLOGICALLY SALIENT TOPIC SUBSETS

No.	Topic	Number of Cases
1	Search & Seizure	25,699
2	Search Warrants	9,702
3	Right to Counsel	9,528
4	Effective Assistance of Counsel	16,221
5	Prisoner Rights	14,049
6	Habeas Corpus	32,707
7	Capital Punishment	3,244
8	Immigration Law	33,657
9	Controlled Substances	17,029
10	Miranda	4,893
11	Guilty Pleas	18,491
12	Civil Rights	44,359
13	Title VII	8,430
14	Affirmative Action	1,495
15	Labor & Employment Law	47,528
16	Discrimination	26,114
17	Employment Law Discrimination	23,124
18	Disability Discrimination	5,142
19	Racial Discrimination	4,492
20	Sex Discrimination	3,332
21	Sexual Harassment	2,182
22	Sexual Orientation	138
23	Workers' Compensation Workers Compensation	11,858
24	Wage & Hour	3,027
25	Collective Bargaining & Labor Relations	14,671
26	Abortion	298
27	Voting Rights	452
28	Campaign	3,424
29	Establishment Clause	147
30	Environmental	10,483
31	Clean Water Act	907

32	Property Rights	1,701
33	Piercing the Corporate Veil	866
34	State Sovereign Immunity	3,110
35	Abrogation of Immunity	487
36	NEPA	4,844
37	NLRB	3,311
38	Campaign Finance	1,176

List of ideologically salient topics and the total number of cases in each topic category. A given opinion typically has more than one topic (including one or more topics that are not ideologically salient).

Table A2

No.	Topic	Number of Cases
1	Civil Rights	44,359
2	Title VII	8,430
3	Affirmative Action	1,495
4	Labor & Employment Law	47,528
5	Discrimination	26,114
6	Employment Law Discrimination	23,124
7	Disability Discrimination	5,142
8	Racial Discrimination	4,492
9	Sex Discrimination	3,332
10	Sexual Harassment	2,182
11	Sexual Orientation	138
12	Abortion	298

List of gender-salient topics and the total number of cases in each topic category. Gender-salient topics are a subset of ideologically salient topics.

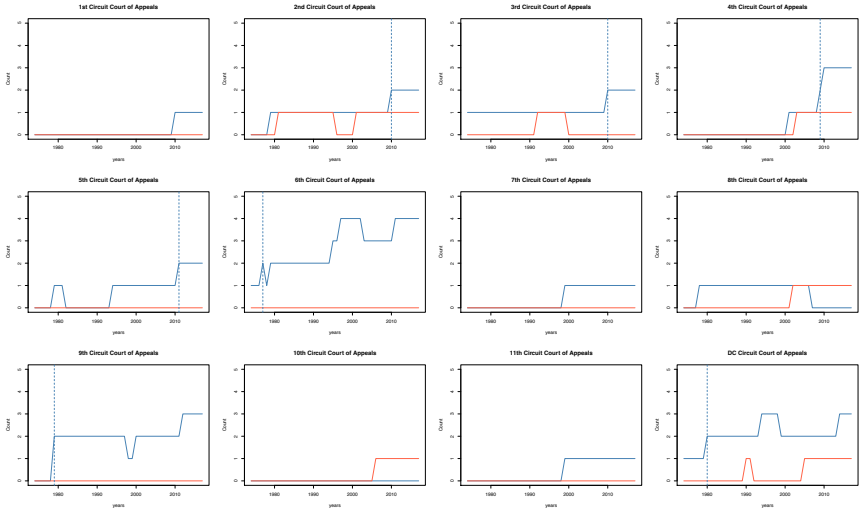
Table A3

No.	Topic	Number of Cases
1	Search & Seizure	25,699
2	Search Warrants	9,702
3	Right to Counsel	9,528
4	Effective Assistance of Counsel	16,221
5	Prisoner's Rights	14,049
6	Habeas Corpus	32,707
7	Capital Punishment	3,244
8	Immigration Law	33,657
9	Controlled Substances	17,029
10	Miranda	4,893
11	Guilty Pleas	18,491
12	Civil Rights	44,359
13	Title VII	8,430
14	Affirmative Action	1,495
15	Discrimination	26,114
16	Employment Law Discrimination	23,124
17	Racial Discrimination	4,492
18	Voting Rights	452

List of race-salient topics and the total number of cases in each topic category. Race-salient topics are a subset of ideologically salient topics.

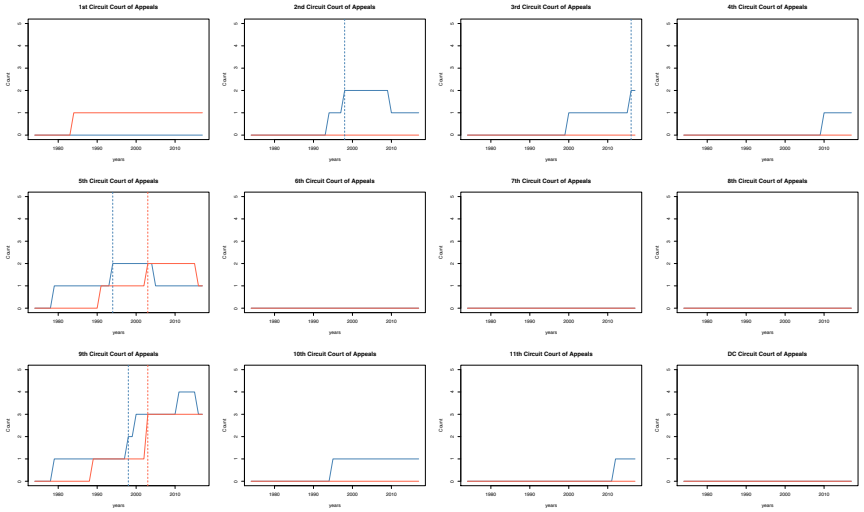
A2. NUMBER OF BLACK, HISPANIC, AND FEMALE JUDGES BY CIRCUIT AND YEAR

Figure A1



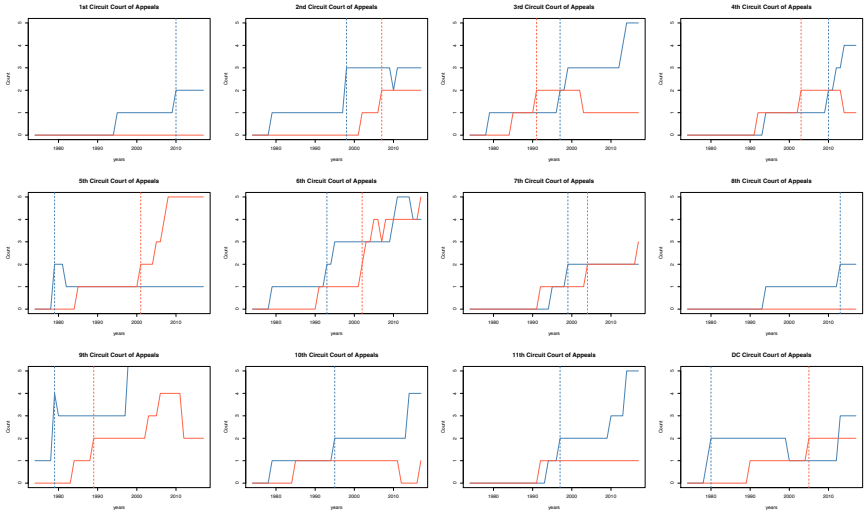
Total number of Black judges for each court and year by partisanship. Red and Blue dotted vertical lines are added to the years that the given circuit has two Black Republican judges and Black Democratic judges respectively. Data from FJC.

Figure A2



Total number of Hispanic judges for each court and year by partisanship. Red and Blue dotted vertical lines are added to the years that the given circuit has two Hispanic Republican judges and Hispanic Democratic judges respectively. Data from FJC.

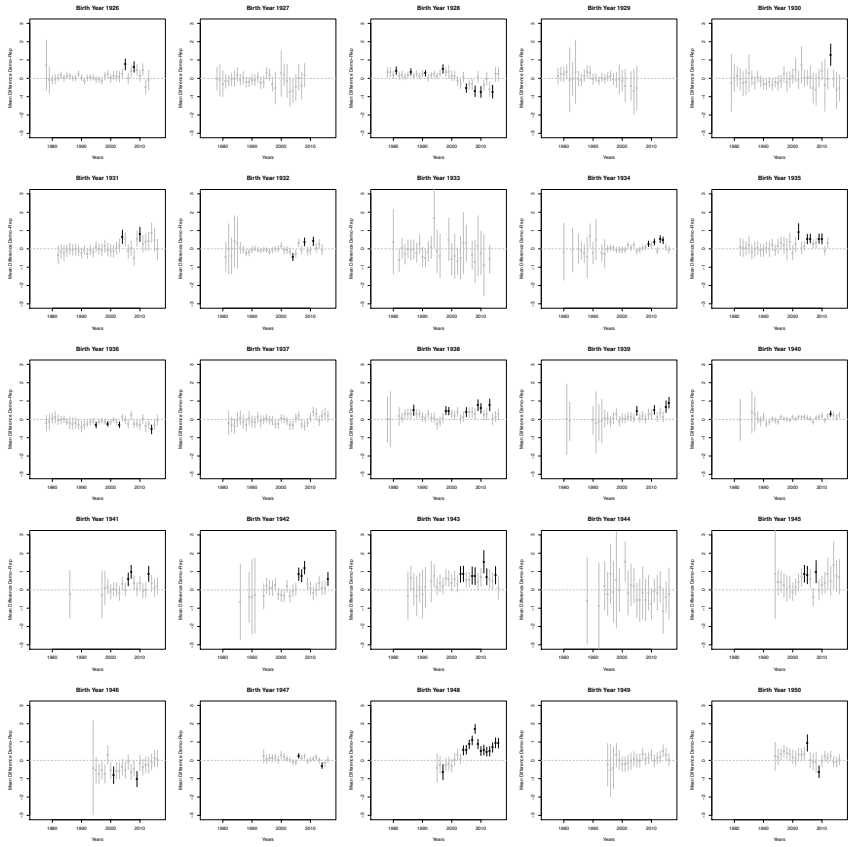
Figure A3



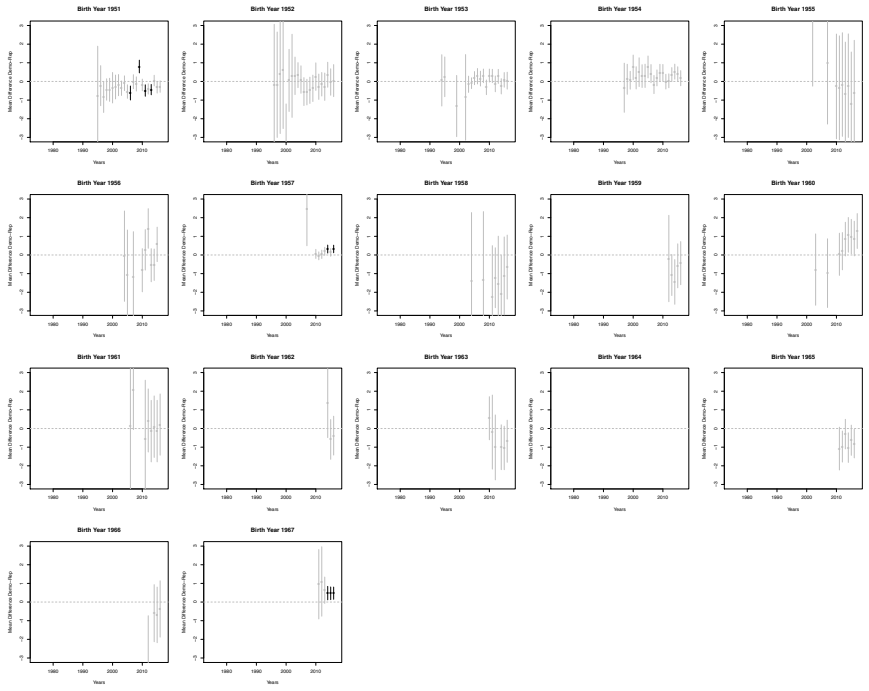
Total number of female judges for each court and year by partisanship. Red and Blue dotted vertical lines are added to the years that the given circuit has two female Republican judges and female Democratic judges respectively. Data from FJC.

A3. AGE OF JUDGES AND THE PARTISAN DIFFERENCES

Figure A4

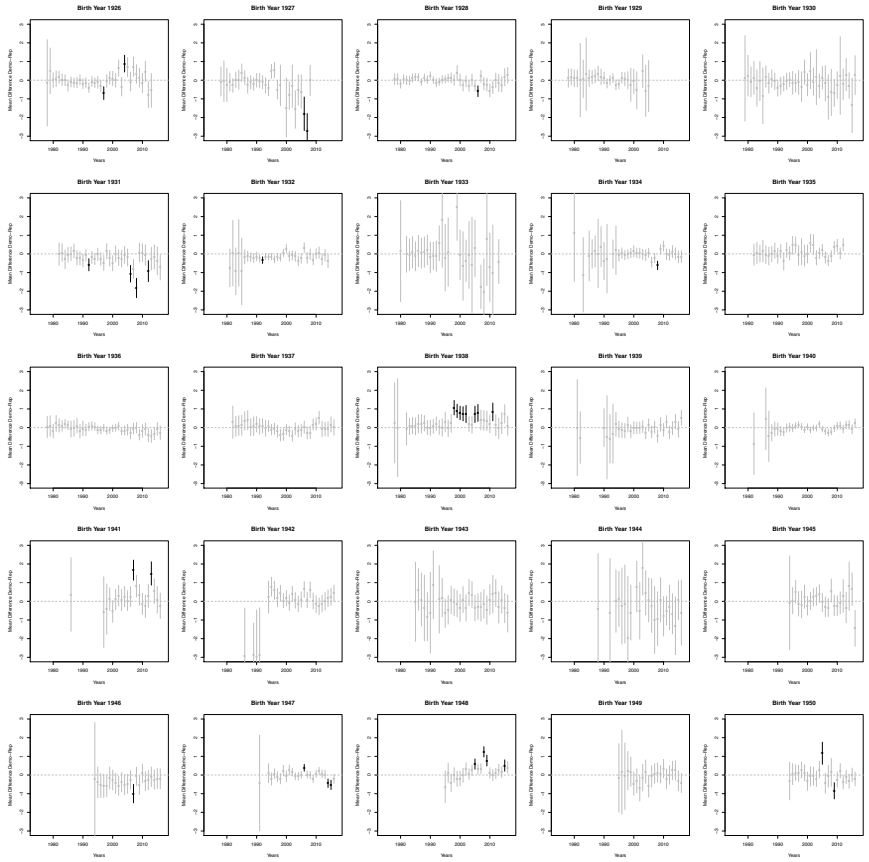


(continues on next page)

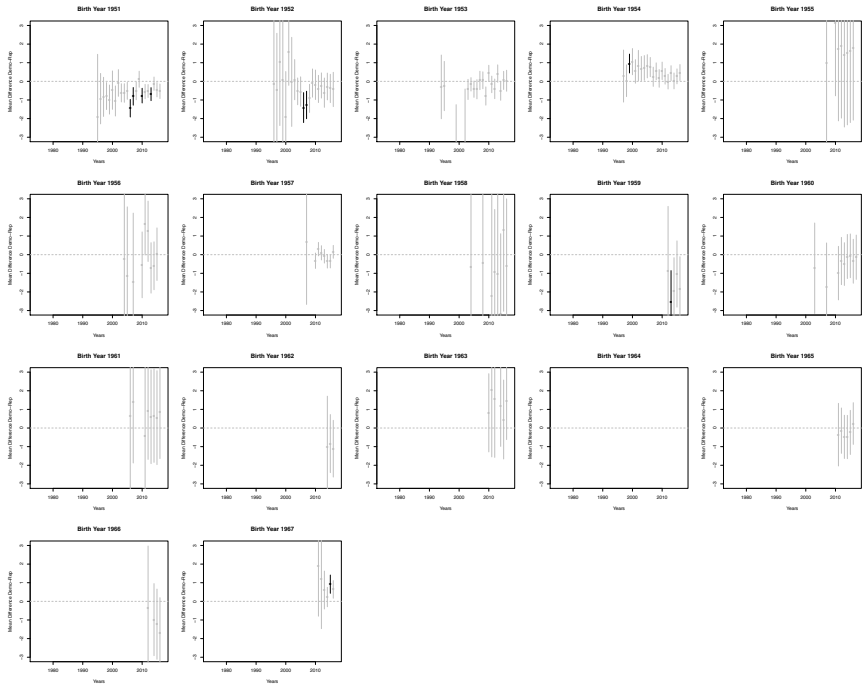


Effect of birth year of appointees on the positive treatment of opinions written by Democratic authors. Each point is a predicted difference between Democratic authors and Republican authors in the propensity to follow opinions by Democratic authors. To obtain predicted values, we fit an OLS regression of the count of positive treatment to Dem-authored opinions on the interaction of author's party and year. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are in black. Statistically insignificant differences are in gray.

Figure A5



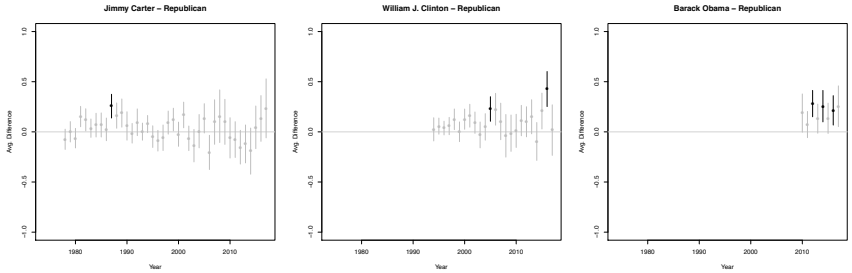
(continues on next page)



Effect of birth year of appointees on the positive treatment of opinions written by Republican authors. Each point is a predicted difference between Democratic authors and Republican authors in the propensity to follow opinions by Republican authors. To obtain predicted values, we fit an OLS regression of the count of positive treatment to Rep-authored opinions on the interaction of author's party and year. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are in black. Statistically insignificant differences are in gray.

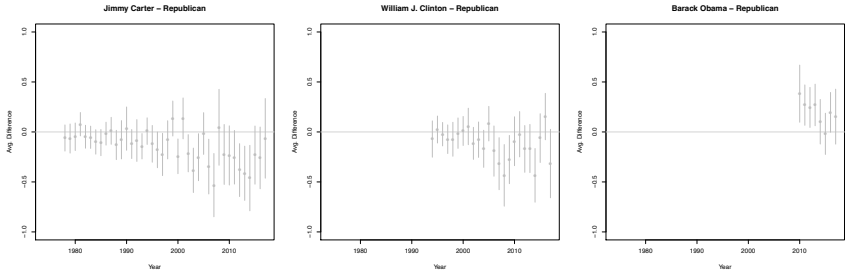
A4. PRESIDENTIAL COHORTS AND PARTISAN DIFFERENCES

Figure A6



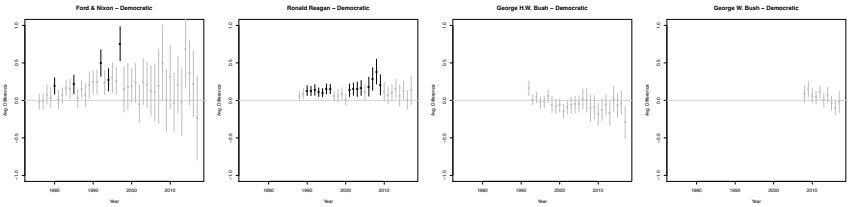
Differences in the average number of positive treatments to opinions authored by Democratic appointees for Democratic presidential cohorts compared to all Republican presidential cohorts. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are in black. Statistically insignificant differences are in gray.

Figure A7



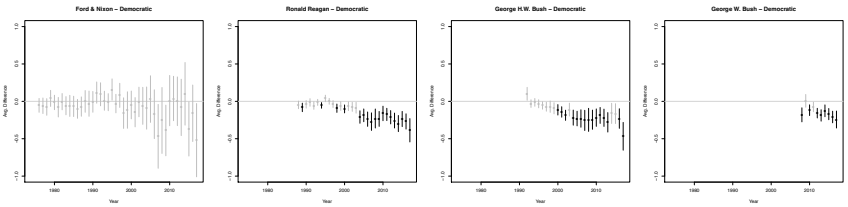
Differences in the average number of positive treatments to opinions authored by Republican appointees for Democratic presidential cohorts compared to all Republican presidential cohorts. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are in black. Statistically insignificant differences are in gray.

Figure A8



Differences in the average number of positive treatments to opinions authored by Republican appointees for Republican presidential cohorts compared to all Democratic presidential cohorts. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are in black. Statistically insignificant differences are in gray.

Figure A9



Differences in the average number of positive treatments to opinions authored by Democratic appointees for Republican presidential cohorts compared to all Democratic presidential cohorts. Differences that are statistically significant at the 0.05 level after adjusting for multiple testing are in black. Statistically insignificant differences are in gray.

"Any"

James J. Brudney & Ethan J. Leib[*]

Our statute books use the word "any" ubiquitously in coverage and exclusion provisions. As any reader of the Supreme Court's statutory interpretation docket would know, a large number of cases turn on the contested application of this so-called universal quantifier. It is hard to make sense of the jurisprudence of "any." And any effort to offer a unified approach—knowing precisely when its scope is expansive (along the "literal-meaning" lines of "every" and "all") or confining (having a contained domain related to properties provided by contextual cues)—is likely to fail. This Article examines legislative drafting manuals, surveys centuries of Court decisions, and conducts in-depth pairwise comparisons of "any" cases to show the word's flexible set of uses in its multiple statutory guises. After evaluating evidence of the variability of "any," we recommend a new approach, a form of an "any" canon. We encourage adjudicators to appreciate the complexity of "any" more systematically and to consult a full range of sources—as even full-throated textualists have authorized from time to time—offering the relevant larger context judges will need to ascertain the scope of "any" in any given statutory scheme.

* James J. Brudney is the Joseph Crowley Chair in Labor and Employment Law at Fordham Law School. Ethan J. Leib is the John D. Calamari Distinguished Professor of Law at Fordham Law School. Thanks to Aaron-Andrew Bruhl, Jesse Cross, Bill Eskridge, Anita Krishnakumar, James Macleod, Vic Nourse, Ed Rubin, Brian Slocum, Larry Solum, and participants in an AALS panel discussion, the Legislation Roundtable at Georgetown Law Center, and a Fordham Faculty Retreat for valuable comments and suggestions. Thanks also to Brennan Corriston and Abigail Tubin for excellent research assistance, and to Fordham Law School for generous financial support.

CONTENTS

INTRODUCTION

Anyone who knows anything about statutory interpretation opinions at the Supreme Court knows that many of them seem to turn on the word "any." We concede that the prior sentence is anything but precise; at the same time, it is not easy to tell how a court—or a legislature—will choose to make use of this word in any given context.

The word "any" seems to be an essential element of statutory drafting.[1] Laws of general application are enacted with an

1. Some of what we will say about "any" could apply to other universal quantifiers like "every" or "all," which also populate our statutes. But in our respective decades of teaching and writing about statutory interpretation, it is "any" that stands out as the universal quantifier most often relevant with respect to statutory drafting and litigation. To take two notable drafting examples, the twenty-nine-page Civil Rights Act of 1964 (CRA) includes 302 instances of "any," more than ten per page. Civil Rights Act of 1964, Pub. L. No. 88-352, 78 Stat. 241. And the 906-page Affordable Care Act of 2010 (ACA) contains 1,774 uses of "any," roughly two per page. Affordable Care Act of 2010, Pub. L. No. 111-148, 124 Stat. 119. This relative difference in frequency for two prominent "super-statutes" warrants

understanding that they will extend for years into the future and are likely to embrace unforeseen settings. Accordingly, legislatures must be attentive to the scope of what actions or individuals they are covering under their laws, and also what or whom they are excluding from coverage. The use of "any" as an adjective modifying a noun or series of nouns is one crucial way in which legislatures can establish that laws apply to, or exempt from coverage, certain groups of individuals or entities.

Congress or state legislatures may use "any" to indicate potentially expansive statutory coverage when establishing a basic norm of criminal or civil conduct.[2] They also may use "any" to reflect a broad exception to a basic statutory norm.[3] In these and other examples, "any," as an adjective followed by a single noun or series of nouns, essentially means "every"—as in "any child knows that" or "any pharmacist will tell you the same thing."

On the other hand, "any" as a modifier may convey a narrower scope, akin to "some," without referring to or implying a quantity. Examples from ordinary usage include "do you have any money?"

---

further attention than we can provide here. One possible explanation is that the CRA, declaring broad normative principles, reflects a heavy dose of prohibitions and protections that result in more extensive use of "any" within basic qualifying sections (compare the federal and state antidiscrimination examples *infra* note 2). By contrast, the ACA focuses more on technical operational guidelines for agencies, both federal (HHS, Labor, IRS) and state, and deals less with prescriptive inclusions or exclusions.

The prominence of "any" in Supreme Court litigation over many decades is the focus of our survey review in Part II and our deeper dive on eight decisions in Part III.

2. *See, e.g.*, 18 U.S.C. § 1961(4) (defining "enterprise" to include "any individual, partnership, corporation, association, or other legal entity, and any union or group of individuals associated in fact although not a legal entity" under the Racketeer Influenced and Corrupt Organization Act (RICO) title of 1970 Organized Crime Control Act); 29 U.S.C. § 633a(a) (specifying that personnel actions affecting employees or applicants at least forty years of age "shall be made free from any discrimination based on age" under the Age Discrimination in Employment Act (ADEA)); OR. REV. STAT. § 659A.142(4) (2021) ("It is an unlawful practice for any place of public accommodation, resort or amusement as defined in ORS 659A.400, or any person acting on behalf of such place, to make any distinction, discrimination or restriction because a customer or patron is an individual with a disability.").

3. *See, e.g.*, 9 U.S.C. § 1 ("[N]othing herein contained [within the Federal Arbitration Act] shall apply to contracts of employment of seamen, railroad employees, or any other class of workers engaged in foreign or interstate commerce."); 28 U.S.C. § 2680(j) (providing that the waiver of sovereign immunity in the Federal Tort Claims Act (FTCA) does not apply to "[a]ny claim arising out of the combatant activities of the military or naval forces, or the Coast Guard, during time of war"); NEB. REV. STAT. § 13-910(7) (2019) (Nebraska Political Subdivisions Tort Claims Act) (exempting from waiver of sovereign immunity "[a]ny claim arising out of assault, battery, false arrest, false imprisonment, malicious prosecution, abuse of process, libel, slander, misrepresentation, deceit, or interference with contract rights").

or "I would appreciate any help you can provide." Courts routinely decide that Congress or state legislatures have used "any" in a limiting sense, indicative of less-than-universal statutory coverage,[4] or a less-than-expansive statutory exception to coverage.[5] Yet courts have expressed frustration at the ambiguous meaning of "any" in a criminal statutory setting.[6]

That Congress or state legislatures may draft expansively or narrowly is hardly news. And there are myriad decisions by the Court since the 1970s that showcase serious disagreements as to what Congress has said or meant when the word "any" appears in various statutory and factual settings. Moreover, depending on a judge's preferences for the scope of "any" in a particular case, string cites are ready-to-hand at the Supreme Court for confining[7] and expansive[8] constructions. Yet the Court—ballasted by Justice

4. *See* Yates v. United States, 574 U.S. 528 (2015) (discussing coverage of "any . . . tangible object" under the Sarbanes-Oxley Act to cover records and documents, but not fish); Home Depot U.S.A, Inc. v. Jackson, 139 S. Ct. 1743 (2019) (deciding that the right of removal for "any defendant" under the Class Action Fairness Act does not include a third-party counterclaim defendant); People v. Davis, 766 N.E.2d 641, 646 (Ill. 2002) (deciding that state armed violence statute did not apply to BB gun because it was not "any other deadly or dangerous weapon or instrument of like nature").

5. *See* Dolan v. U.S. Postal Serv., 546 U.S. 481, 494 (2006) (evaluating the exclusion of "*any* claim arising out of . . . 'negligent transmission'" from the sovereign immunity waiver in the FTCA, and limiting it to only certain forms of negligent transmission); Univ. of Tex. at Arlington v. Williams, 459 S.W.3d 48, 51 (Tex. 2015) (holding that statutory limit on landowners' liability when opening premises for public recreational use, including inter alia for hunting, fishing, swimming, birdwatching and "any other activity associated with enjoying nature or the outdoors[,]" does not limit liability for injury to spectator watching a soccer game).

6. *See* State v. Fourth Jud. Dist. Ct. of Nev., 481 P.3d 848 (Nev. 2021). The Nevada statute there prohibited a felon from possessing "any firearm"; the defendant, a convicted felon, possessed five firearms at a single time and place. The Nevada Supreme Court noted that other state criminal laws similarly use "any" to help define prohibited conduct; it concluded that "ambiguity arises because '"[t]he word "'any"' has multiple, conflicting definitions, including (1) one; (2) one, some, or all regardless of quantity; (3) great, unmeasured, or unlimited in amount; (4) one or more; and (5) all.'" *Id.* at 850 (internal citation omitted). After determining that other tools of interpretation did not resolve the relevant textual ambiguity, the court applied the rule of lenity, holding there was only a single violation of the felon-in-possession statute.

7. *See, e.g.*, Small v. United States, 544 U.S. 385, 388 (2005) (citing Nixon v. Mo. Mun. League, 541 U.S. 125, 132 (2004); United States v. Alvarez-Sanchez, 511 U.S. 350, 357 (1994); Middlesex Cnty. Sewerage Auth. v. Nat'l Sea Clammers Ass'n, 453 U.S. 1, 15–16 (1981); Flora v. United States, 362 U.S. 145, 149 (1960)).

8. *See, e.g.*, *Yates*, 574 U.S. at 555 (Kagan, J., dissenting) (citing Dep't of Hous. and Urb. Dev. v. Rucker, 535 U.S. 125, 131 (2002); Republic of Iraq v. Beaty, 556 U.S. 848, 856 (2009); Ali v. Fed. Bureau of Prisons, 552 U.S. 214, 219–20 (2008)).

Kagan's once-approving observation that "we are all textualists now"[9]—has increasingly maintained that deciphering "any" in congressional work-product is or should be a reasonably straightforward textualist enterprise. To that end, the justices have taken to referring to the word's "ordinary" meaning,[10] and looking up "any" in the dictionary as confirmation of that concept,[11] including in the Court's most recent term.[12] At the same time, the Court's cases reveal tensions between a literal, dictionary-driven modality of defining "any" and an alternative way of finding "ordinary" meaning that is shaped or restrained by statutory context, more broadly understood.[13]

This Article unfolds in four parts. First, we explore drafting manuals from Congress and state legislatures to get some sense of what legislatures think they are doing when including the word "any" in statutes. In Part II, we survey the Supreme Court's many decisions invoking or explaining "any" in a statute to identify trends and furnish initial analysis about the Court's disparate efforts to figure out what "any" means. We show that "any" can transcend traditional ideological alignments: Clashes feature conservative justices sharply disagreeing with one another and liberal justices at odds with other liberals. Even the Court's liberal wing, which is more willing to look at extrinsic sources in some kinds of cases, goes oddly literalist at times when it comes to "any." Part III then explores in detail four pairs of cases. These deeper dives reveal that a hyperfocus on text is unlikely to provide a useful "any" jurisprudence. Instead, judges must strive to better appreciate the statutory setting in which "any" is situated. This in turn

9. *See* Harvard Law School, *The 2015 Scalia Lecture | A Dialogue with Justice Elena Kagan on the Reading of Statutes*, YOUTUBE at 08:28 (Nov. 25, 2015), https://youtu.be/dpEtszFT0Tg [https://perma.cc/6HMD727M]. *But see* West Virginia v. EPA, 142 S. Ct. 2587, 2641 (2022) (Kagan, J., dissenting) ("The current Court is textualist only when being so suits it.").

10. *See* United States v. Gonzalez, 520 U.S. 1, 5 (1997) (Thomas, J.) (invoking the "ordinary sense" of "in any court"); *Yates*, 574 U.S at 556 (Kagan, J., dissenting) (invoking the "ordinary meaning" of "any physical object").

11. *Gonzalez*, 520 U.S. at 5 (Thomas, J.) (citing WEBSTER'S THIRD NEW INTERNATIONAL DICTIONARY 97 (1976)); *Yates*, 574 U.S. at 555 (Kagan, J., dissenting) (citing WEBSTER'S THIRD NEW INTERNATIONAL DICTIONARY 97 (2002)).

12. Patel v. Garland, 142 S. Ct. 1614, 1622 (2022) (Barrett, J.) (citing WEBSTER'S THIRD NEW INTERNATIONAL DICTIONARY 97 (1993)).

13. *See, e.g.*, Ali v. Fed. Bureau of Prisons, 552 U.S. 214 (2008), discussed *infra* Part III.A; Small v. United States, 544 U.S. 385 (2005), discussed *infra* Part III.C.

requires greater attention to how linguistic and substantive canons—applied in light of statutory purposes and plans—help inform the function and scope of "any" in particular cases. Part IV concludes by linking our analysis and arguments to some recent scholarship on "textual isolationism,"[14] contextual cherry-picking,[15] and ordinary meaning,[16] all of which might be usefully brought to bear on the Court's continuing problem with interpreting the word "any" in statutes.

Unpacking the Court's variable approaches to "any" illustrates a persistent indeterminacy of textualism and the central role of contextual framing for the interpretive process. Our bottom line is that it is very hard to learn the ordinary meaning of the word "any" in a statute without a rich understanding of context. Both textualist and purposivist justices should be willing to look at a range of intrinsic and extrinsic sources in order to grasp what Congress was up to when it used the language of "any." Even the more devout textualists must stop looking up "any" in the dictionary or just italicizing it, and instead acknowledge what might be an "any" canon: "any" is an invitation to look for signals about its scope in a wide set of sources that include linguistic canons, substantive canons, and statutory plans or schemes. We show that textualists also have allowed themselves this critical context from time to time and argue that they should approach this ubiquitous word with modesty and humility.

I. LEGISLATIVE DRAFTING AND THE USES OF "ANY"

Laws of general application will always require some attention to clarifying a law's scope. In this regard, "any" is known by linguists and philosophers as a "universal quantifier."[17] The literally expansive domain of a universal quantifier, however, is not always

14. *See* Victoria Nourse, *Picking and Choosing Text: Lessons for Statutory Interpretation from the Philosophy of Language*, 69 FLA. L. REV. 1409 (2017).

15. *See* William N. Eskridge, Jr. & Victoria F. Nourse, *Textual Gerrymandering: The Eclipse of Republican Government in an Era of Statutory Populism*, 96 N.Y.U. L. REV. 1718 (2022).

16. *See* Kevin Tobia, Brian G. Slocum & Victoria Nourse, *Ordinary Meaning and Ordinary People*, 117 U. PA. L. REV. 365 (2023).

17. *See generally* BRIAN G. SLOCUM, ORDINARY MEANING: A THEORY OF THE MOST FUNDAMENTAL PRINCIPLE OF LEGAL INTERPRETATION (2015). Slocum discusses the problem of universal quantifiers and how literal meaning and ordinary meaning can pull apart in linguistic usage of them—with courts legally construing them often to have literal meanings rather than ordinary meanings. *E.g.*, *id.* at 26, 32–33, 153–67.

what is understood by an ordinary speaker and user of language, who may identify or implicitly rely on ways to restrict or adjust its scope.[18] Similarly, legal drafters sometimes choose to pursue a usage that confines the domain of the universal quantifier and other times prefer to incorporate an approach that is unabashedly expansive. The use of "any" can establish that laws (i) apply broadly and without limitation; (ii) apply only to certain identified groups of individuals or entities, with the implication of not extending beyond those groups; (iii) apply to a presumptively circumscribed group limited by certain properties or characteristics while also contemplating broad and potentially unanticipated group members that arise in the future; or (iv) specify omissions from coverage for certain identified groups in each of the three ways coverage can be delineated—that is, without limitation, based on a limited enumeration, or reflecting an enumeration that is illustrative and broad. Part of the mystery of "any" is that it can do all of this varied work, sometimes in one statute, both for coverage provisions and exemption provisions. When courts are presented with interpretive disputes, statutes can be littered with the word "any," each instance actually serving a different expositional function for judges to decode.

Although legislative drafters regularly include "any" as part of statutory provisions, relatively little specific guidance exists in federal and state drafting manuals.[19] A short discussion in one federal manual offers general advice regarding the utility of the word. As an adjective modifying a singular noun, "any" may

18. *See generally* Jason Stanley & Zoltán Gendler Szabó, *On Quantifier Domain Restriction*, 15 MIND & LANG. 219 (2000) (investigating the special case of context dependence—how context helps us understand utterances—that is quantifier domain restriction, in which the interpretation of words necessitates limits in scope).

19. The manuals tend to focus heavily on organization and structure of bills, with detailed attention to aspects such as effective date, amendments, severability, appropriations, resolutions, and orders; as well as on general rules of grammar, punctuation, and citation to state law materials. The U.S. HOUSE LEGISLATIVE COUNSEL'S MANUAL ON DRAFTING STYLE (1995) runs sixty-four pages [hereinafter 1995 HOUSE LEGIS. MANUAL]. State drafters' manuals often run much longer. *See, e.g.*, MAINE LEGISLATIVE DRAFTING MANUAL (2016) (234 pages); DELAWARE LEGISLATIVE DRAFTING MANUAL (2019) (210 pages); ILLINOIS BILL DRAFTING MANUAL (2012) (264 pages). More access to drafting manuals is available at *Online Drafting Manuals*, NAT'L CONF. OF STATE LEGISLATORS, https://www.ncsl.org/legislators-staff/legislative-staff/research-editorial-legal-and-committee-staff/bill-drafting-manuals.aspx (last updated May 3, 2022).

illuminate breadth of coverage by reducing certain ambiguities,[20] making clear that a statutory provision covers all possible individuals, entities, substances, or actions within the identified universe.[21] Relatedly, "any" may be used at times to convey special emphasis regarding the breadth of a rule.[22]

In addition, drafters may use "any" as a concise way to address uncertainty regarding whether a stated condition will actually occur, including when the condition itself is neither mandated nor perhaps even likely. For example, a law might provide for the EEOC to update a certain statutory test governing retirement benefits "in accordance with any standards established by the Commission."[23] Under this language, the agency may decide not to promulgate such standards, and it is not required to do so. Yet if the word "any" were removed, this would imply that the agency is expected to issue standards on the statutory retirement benefits test.[24]

Beyond descriptions of positive uses for "any," drafting manuals may feature words of caution regarding ambiguities the word can create. Some state manuals discourage the use of "any" for basic grammatical reasons, preferring a singular subject,[25] or

20. *See* 1995 HOUSE LEGIS. MANUAL 61 ("'Any employee who . . .' works the same as 'Employees who . . .' yet it avoids any misreading that (1) an implicit precondition exists that 2 [or more] employees must be involved before either gets covered, or (2) the statement only applies to a group of employees, as such.").

21. *See, e.g.*, 42 U.S.C. § 7602(g) ("The term 'air pollutant' means any air pollution agent or combination of such agents, including any physical, chemical, biological, radioactive (including source material, special nuclear material, and byproduct material) substance or matter which is emitted into or otherwise enters the ambient air. Such term includes any precursors to the formation of any air pollutant, to the extent the Administrator has identified such precursor or precursors for the particular purpose for which the term '"air pollutant'" is used.") (relied on in Massachusetts v. EPA, 549 U.S. 407 (2007)).

22. *See* 1995 HOUSE LEGIS. MANUAL 62.

23. The example is borrowed loosely from 29 U.S.C. § 631(c)(2), which deals with possible adjustment of a retirement benefit test under the Age Discrimination in Employment Act to provide for computation if not based on a straight life annuity. We are grateful to Jesse Cross for suggesting how use of "any" in the drafting process may help navigate such conditional situations.

24. If "any" is included, and should the agency decide to issue the standards, this would trigger the updating of the test.

25. *See* PENNSYLVANIA CODE & BULLETIN STYLE MANUAL § 9.2 (5th ed.) (use "A nursing home requesting"; do not use "Any nursing home requesting"); LEG. COUNS. COMM., OREGON BILL DRAFTING MANUAL ch. 4.6 (18th ed. 2018) ("Simple words such as '"a,'" '"an'" or '"the'" nearly always can be used instead of '"any[]'" . . . with an attendant gain in clarity.").

warn against "common legalisms that are often unclear and nearly always unnecessary."[26]

State drafters' manuals also focus on the relevance of context to the use of "any." One manual cites to a U.S. Supreme Court decision when observing that, "[r]ead naturally, the word 'any' has an expansive meaning, that is, 'one or some indiscriminately of whatever kind.'"[27] At the same time, manuals express concern about settings in which "any" begins or ends an enumeration of particulars that includes general words or phrases in association with specific words. Illinois warns drafters, "Be careful when adding to a list of specific items that ends with a catchall item such as 'any other information required by the Director.'"[28] Oregon counsels that when a provision is to apply to an entire class, "it is generally safer if the class is named in general terms . . . even when the particulars [named] would be preceded or followed by general language. It is virtually impossible to make an enumeration all-inclusive, and omission may be construed as implying deliberate exclusion."[29] And Delaware observes that when used at the start of a non-exclusive enumeration provision, "any" can give rise to reasonable inferences of broad coverage unless words are added to negate such inferences.[30]

---

26. *See* THE OFF. OF THE REVISOR OF STATUTES, MINNESOTA REVISOR'S MANUAL ch. 8.25 (2013 ed.) (recommending use of "a, an, the" and not "all, each, every, some," and presumably "any" for the same reason). The most recent House Legislative Manual also discourages the use of "any," indicating that the preferred style is "a" or "an" and that "any" should be used "only when necessary for special emphasis." THE OFF. OF THE LEG. COUNS., U.S. HOUSE LEGISLATIVE COUNSEL'S MANUAL ON DRAFTING STYLE 47 (Dec. 2022). Were drafters themselves to heed the calls—try ridding yourself of the use of "any" for a day to appreciate the challenge of being perfectly disciplined—it is likely true that interpretive disputes would be displaced onto other words with unclear scope. We would likely recommend contextual analysis then, too, but "any" is sufficiently at the center of so many statutory interpretation decisions that drawing attention to the issues here may well have trickle-down effects on interpretation whenever a court has to establish the domain or scope of words with a range of possible uses.

27. OREGON BILL DRAFTING MANUAL, *supra* note 25, at ch. 4.6 (quoting United States v. Gonzalez, 520 U.S. 1, 5 (1997)).

28. THE LEG. REFERENCE BUREAU, ILLINOIS BILL DRAFTING MANUAL § 15.20, at 23 (10th ed. 2012).

29. OREGON BILL DRAFTING MANUAL, *supra* note 25, at ch. 4.10(c).

30. *See* LEG. COUNCIL DIV. OF RSCH., DELAWARE LEGISLATIVE DRAFTING MANUAL, r. 28(e) at 115 (citing to a public records code provision stating that "'Person' includes an individual, corporation, business trust, estate trust, partnership, association, joint venture, . . . or any other legal or commercial entity. [The term] *does not include* a government;

In light of such concerns about the effects of using general words like "any" in association with particular words, it is not surprising that some manuals invoke *ejusdem generis*, a familiar canon of construction related to word association. The Oregon Manual has a section on *ejusdem generis* in this regard, explaining that it is "based on the reasonable assumption that a drafter will not enumerate items if the drafter intends general words to have their unrestricted meaning," and then illustrating with an example from a state supreme court case.[31]

The Delaware Manual provides additional guidance on how to draft enumeration provisions. It addresses *provisions that are exclusive* (use only the specific terms meant to be included; omit "includes"; and omit all general terms such as terms preceded by "any other"); *provisions that are nonexclusive or exemplary* (use "include" to precede the list; use both specific and general terms in constructing the list); and *provisions that are nonexclusive with some limitation* (use specific and general terms "of the same type or nature as each other"; the general term "must be one that clearly belongs to the same class as the specific terms").[32]

Delaware explains that guidance on drafting exclusive provisions follows *expressio unius*, while guidance for drafting nonexclusive provisions with some limitation comports with *ejusdem generis*.[33] Regarding this latter guidance, the manual's Comment invokes a state court decision for emphasis, noting that "if the Legislature had intended the general word to be used in an unrestricted sense, no mention would have been made at all of the particular classes" and that the "words 'other' or 'any other' following an enumeration of particular classes are to be read as meaning 'other such like' and include only words of like kind or

---

a governmental subdivision, agency, or instrumentality; or a public corporation[,]" and observing that "[w]ithout the negating sentence in this example, one could reasonably infer that a governmental body is within the scope of the definition as 'any other legal or commercial entity.'").

31. *See* OREGON BILL DRAFTING MANUAL, *supra* note 25, at ch. 4.10(b). The manual recounts specifics of the case, quoting *State v. Brantley*, 201 Or. 637, 645–46 (1954). "[A] statute that applied to *any* forged "'record, writing, instrument or matter whatever'" was held not to apply to [a forged] certificate of nomination for [public office] candidacy . . . , because . . . [t]he words "'or matter whatever'" were limited . . . by the preceding enumeration of particulars[,]" and the certificate was not a record, writing, or instrument as those terms were defined by law.

32. *See* DELAWARE LEGISLATIVE DRAFTING MANUAL, *supra* note 30, at r. 29C (c), (d), (e).

33. *See id.* at 120–21 (Comment on Rule 29C).

character."[34] Whether the Delaware legislature actually drafts according to the parameters in the drafting manual is a different question, but the manual at least underscores that certain familiar canons of word association can be understood as consistent with expressed legislative intent.[35]

Many manuals, unlike those just discussed, lack in-depth or nuanced analysis of the benefits and costs associated with using "any" in statutory text. Nonetheless, manuals in a number of states and the U.S. House pay attention to how "any" may assume constructively clear or concerningly cloudy usages based on surrounding context. This bottom-line insight from legislative drafters was echoed by a Supreme Court Justice in a relatively recent case, which we explore more carefully in Part III:[36] "it is context, not a dictionary, that sets the boundaries of time, place, and circumstances within which words such as 'any' will apply."[37]

Given that legislatures use "any" in both expansive and confining ways, and that they may add neighboring words to help reinforce or cabin the scope of this term, is there any lesson to be gleaned from the guidance transmitted by legislative drafters? One might borrow here from Richard Posner's counsel for judges to "try to put [themselves] in the shoes of the enacting legislators . . . ."[38] And insofar as this is not possible, Posner's advice suggests that the scope of a word like "any" should be based on "what attribution of meaning . . . will yield the most reasonable result . . . bearing in mind . . . that it is [the legislators'] conception of reasonableness, to the extent known, rather than the judge's, that should guide decision."[39] To be sure, strict textualists of the ordinary meaning subclass, who focus on citizen beholders of the statute rather than

34. *See id.* (quoting Bigger v. Unemployment Comp. Comm'n, 46 A.2d 137, 141 (Del. Super. Ct. 1946), *aff'd*, 53 A.2d 761 (Del. 1947)).

35. For some evidence that federal legislative staff drafts with respect for the potential applications of *ejusdem generis* and *noscitur a sociis*, even if drafters aren't aware of the Latin canons by name, see Abbe R. Gluck & Lisa Schultz Bressman, *Statutory Interpretation from the Inside – An Empirical Study of Congressional Drafting, Delegation, and the Canons: Part I*, 65 STAN. L. REV. 901, 933 (2013).

36. *See infra* Part III.A.

37. Ali v. Fed. Bureau of Prisons, 552 U.S. 214, 244 (2008) (Breyer, J., dissenting). *See also* United States v. Palmer, 16 U.S. (3 Wheat.) 610, 631 (1818) (Marshall, C.J.) ("[G]eneral words," such as the word "'any'," must "be limited" in their application "to those objects to which the legislature intended to apply them.").

38. RICHARD A. POSNER, THE FEDERAL COURTS: CRISIS AND REFORM 286–87 (1985).

39. *Id.* at 287.

average legislators passing the statute, might want a different benchmark.[40] But as we have been urging and will continue to show, "any" doesn't have just one ordinary meaning—and it is perilous to presume the literal expansive scope of the universal quantifier always prevails. In the two Parts to follow, we examine what the Supreme Court has done over the years with this term—and to what extent it is Congress, a conception of a reasonable Congress, or something else that has been driving results.

II. A SURVEY OF "ANY" IN THE SUPREME COURT

It is not surprising that Congress and state legislatures may draft either expansively or narrowly when they invoke the word "any." At the same time, decisions by the Supreme Court over many decades showcase disagreements as to what Congress has said or meant when using "any"—whether to cover individuals or actions, or to exclude them. Although these disagreements have become more linguistically focused in the Roberts Court, the results are no more predictable.

A. From Context to Text

The Court's evolution in its approach to "any" is part of a more general shift from routine reliance on purpose and legislative history to reliance on adjacent language and statutory structure, or simply ignoring statutory context in favor of literal meaning presented by dictionary definitions. In an early and famous case construing "any," the Court recognized that the plain meaning of an immigration statute that prohibited "in any way" assisting or encouraging the importation of "any foreigner" under contract "to perform labor or service of any kind" did, in fact, include within its coverage a pastor for New York's leading Protestant church.[41] Still, the Court concluded that the "spirit" of the statute, as understood through its purpose and legislative history, trumped this unambiguously expansive plain meaning.[42] Other decisions in the

40. *But see* James J. Brudney & Lawrence Baum, *Oasis or Mirage: The Supreme Court's Thirst for Dictionaries in the Rehnquist and Roberts Eras*, 55 WM. & MARY L. REV. 483, 533–35 (2013) (reporting that from the 1989 to 2010 terms, strict textualists Scalia and Thomas often relied on dictionaries close to an enactment date, on dictionaries close to the date a dispute arose, and on dictionaries published close to neither date of enactment nor date of case filing).

41. *See* Holy Trinity Church v. United States, 143 U.S. 457, 458 (1892).

42. *See id.* at 458–64.

middle and latter part of the last century established the breadth of statutory coverage for "any" by first addressing purpose and legislative history, and then invoking the text to reinforce that reading.[43] Alternatively, the Court has paid homage to text at the start of a decision but then later dwelled on purpose and history as underlying or reinforcing the result.[44]

Even in the early 1990s, when the Court was faced with interpreting "any" as part of a key phrase in statutory text, it recognized the applicability or relevance of arguments based on congressional purpose.[45] More recently, though, reflecting the Court's larger interpretive approach, decisions construing "any" have often focused primarily on what is deemed the plain meaning of that word, as defined expansively in dictionaries, with support gleaned from string citations to earlier expansive plain meaning decisions.[46]

---

43. *See, e.g.*, United States v. Rosenwasser, 323 U.S. 360, 361–63 (1945) (construing "any of his employees" in a Fair Labor Standards Act overtime provision, and invoking Senate committee report and floor statement of leading sponsor); Shea v. Vialpando, 416 U.S. 251, 258–61 (1974) (construing "any expenses" under Social Security Act provisions as they evolved over several Congresses, and relying on this statutory history as well as committee reports).

44. *See, e.g.*, NAACP v. New York, 413 U.S. 345, 353–55 (1973) (broadly interpreting "any appeal" within a section of Voting Rights Act of 1965); United States v. Turkette, 452 U.S. 576, 580–81, 587–93 (1981) (broadly construing "any enterprise" within RICO). *See also* Harrison v. PPG Indus., Inc., 446 U.S. 578, 587–94 (1980) (discussed *infra* Part III.B) (broadly interpreting "any other final action of the Administrator" under Clean Air Act Amendments of 1977 after considering both *ejusdem generis* and legislative history).

45. *See, e.g.*, Reves v. Ernst & Young, 494 U.S. 56, 61–64, 67–70 (1990) (discussed *infra* Part III.D) (construing "any note" within definition of "security" in 1934 Securities Exchange Act based on congressional purpose to regulate all instruments sold as investments); MCI Telecomms. Corp. v. Am. Tel. & Tel. Co., 512 U.S. 218, 225–27, 229–31 (1994) (construing agency's authority to "modify any requirement" under tariff-filing provisions of 1934 Federal Communications Act by invoking plain meaning and also central role of rate-filing requirement to Congress's regulatory purpose).

46. *See, e.g.*, United States v. Gonzalez, 520 U.S. 1, 5 (1997) (relying on dictionary definition to construe provision of criminal code); Ali v. Fed. Bureau of Prisons, 552 U.S. 214, 218–20 (2008) (discussed *infra* Part III.A) (relying on *Gonzalez*—including its dictionary definition—and *Harrison*, 446 U.S. at 587–94 when construing provision of FTCA); Babb v. Wilkie, 140 S. Ct. 1168, 1173 & n.2 (2020) (relying on dictionary definition to construe provision of law regulating age discrimination in federal employment, and citing to *Ali* and *Gonzalez* for support: "We have repeatedly explained that 'the word "any" has an expansive meaning.'"); Patel v. Garland, 142 S. Ct. 1614, 1621–22 (2022) (relying on four dictionary definitions and citing to *Babb* and *Gonzalez* when construing provision of Immigration and Nationality Act).

B. Expansive Versus Confining Interpretations

Yet the Court in recent decades has not consistently construed "any" as having an expansive scope, despite suggestions to the contrary in certain settings.[47] Since 1990, the Court's caselaw is actually fairly evenly divided between expansive and confining constructions of "any." Expansive readings primarily invoke the straightforward textualist mantra of plain or "ordinary" meaning and dictionary definitions.[48] Confining constructions rely instead on various contextual factors, including word association canons such as *ejusdem generis* and *noscitur a sociis*,[49] the whole act canon and related structural considerations;[50] substantive canons such as the presumptions against federal preemption of state laws[51] and against extraterritorial application of legislation;[52] and the meaning of words accompanying "any" within compact textual phrases.[53] The Roberts Court, however, rarely relies openly on purpose—much less legislative history—when seeking to discern the function of "any" in these statutory settings.[54]

47. *See Babb*, 140 S. Ct. at 1173 n.2.

48. *See, e.g.*, *id.*; *Gonzalez*, 520 U.S. at 5; Massachusetts v. EPA, 549 US 497, 528–29 (2007); *Ali*, 552 U.S. at 218–20 (2008); *Patel*, 142 S. Ct. at 1621–22.

49. *See, e.g.*, Dolan v. U.S. Postal Serv., 546 U.S. 481, 486–87 (2006) (discussed *infra* Part III.A); Cir. City Stores, Inc. v. Adams, 532 U.S. 105, 114–15 (2001) (discussed *infra* Part III.B.2); Yates v. United States, 574 U.S. 528, 543–45 (2015). To be fair, the Court on occasion takes these word association canons to be evidence of plain meaning or ordinary meaning—and more generally the textual canons are sometimes used to clear up confessed ambiguities. There is some uncertainty about the role the textual canons serve within the textualism of the moment.

50. *See, e.g.*, Gutierrez v. Ada, 528 U.S. 250, 254–55 (2000); *Yates*, 574 U.S. at 543–45; Home Depot U.S.A., Inc. v. Jackson, 139 S. Ct. 1743, 1749–51 (2019).

51. *See, e.g.*, Nixon v. Mo. Mun. League, 541 U.S. 125, 140–41 (2004) (discussed *infra* Part III.C).

52. *See, e.g.*, Small v. United States, 544 U.S. 385, 388–91 (2005) (discussed *infra* Part III.C).

53. *See, e.g.*, MCI Telecomms. Corp. v. Am. Tel. & Tel. Co., 512 U.S. 218, 225 (1994) (construing "modify any requirement"); Freeman v. Quicken Loans, Inc., 566 U.S. 624, 633–35 (2012) (construing "any portion, split, or percentage"); Nat'l Ass'n of Mfrs. v. Dep't of Def., 138 S. Ct. 617, 628–29 (2018) (construing "any effluent limitation").

54. *But cf. Yates*, 574 U.S. at 532 (invoking purpose of underlying statute at outset of analysis, before relying heavily on various canons and structural factors to construe narrowly a statute criminalizing the destruction of "any record, document, or tangible object"). In Parts III and IV, we do our best to encourage even textualist judges to remain open to purposive readings of "any," as the alternatives are insufficient to produce a reliable or persuasive "any" jurisprudence.

C. Minimal Role of Ideology in Disagreements Among Justices

Decisions presenting disagreement about the application of "any" have involved liberal justices doing intramural battle as well as conservatives facing off against one another.[55] In addition, justices are not always consistent in their methodological approaches; they may pledge allegiance to the plain meaning expansiveness of "any" in some instances yet also author decisions imposing limits on scope.[56] Further, the link between expansive or confining interpretations of "any" and ideologically-oriented results is not terribly instructive. Cases construing "any" in expansive terms have resulted in both liberal and conservative outcomes.[57] The same may be said for cases that have imposed a limiting construction.[58] Finally, while a handful of decisions construing "any" are unanimous,[59] the cases more often involve disagreements as to the scope and significance of "any" in different statutory settings.

D. The Textualist Trump Card

This is now a solidly textualist as well as largely Republican-appointed Supreme Court. Thus, even justices who believe in the importance of statutory purpose and legislative history[60] rarely rely

55. Examples here include *Yates* (Justice Ginsburg in majority and Justice Kagan in dissent); *Ali v. Fed. Bureau of Prisons*, 552 U.S. 214 (2008) (Justice Thomas in majority and Justice Kennedy in dissent); and *Patel v. Garland,* 142 S. Ct. 1614 (2022) (Justice Barrett in majority and Justice Gorsuch in dissent).

56. For example, compare Justice Thomas's expansive positions on "any" in *Ali* and his dissent in *Dolan v. U.S. Postal Serv.*, 546 U.S. 481 (2006), with the constraints he imposed on the reach of "any" in *United States v. Alvarez-Sanchez*, 511 U.S. 350, 357 (1994) and *Home Depot U.S.A., Inc. v. Jackson*, 139 S. Ct. 1743, 1750 (2019).

57. Compare *Babb v. Wilkie*, 140 S. Ct. 1168 (2020) and *Massachusetts v. EPA*, 549 U.S. 497 (2007), which use expansive constructions to deliver "liberal" results, with *United States v. Gonzalez*, 520 U.S. 1 (1997) and *Ali*, which use expansive constructions to render "conservative" results.

58. Compare *Small v. United States*, 544 U.S. 385 (2005), *Dolan*, and *Yates*, which use narrow constructions to secure "liberal" results, with *Cir. City Stores, Inc. v. Adams*, 532 U.S. 105 (2001) and *Alvarez-Sanchez*, 511 U.S., which use narrow constructions to achieve "conservative" results.

59. *See, e.g.*, Nat'l Ass'n of Mfrs. V. Dep't of Def., 138 S. Ct. 617, 628–29 (2018) (limiting construction); *Gonzalez*, 520 U.S. at 5 (expansive construction); Gutierrez v. Ada, 528 U.S. 250 (2000) (limiting construction).

60. Over the past decade, justices who have openly endorsed such an approach have been those appointed by Democratic Presidents: Justices Ginsburg, Breyer, Sotomayor, Kagan, and, perhaps, the recently appointed Justice Jackson.

prominently on those resources when seeking to command a majority for their opinions. That said, in the past several terms, the Court's dominant textualist wing has generated numerous dueling statutory opinions,[61] including on the scope of "any."[62] And while Justices Thomas and Alito in particular have exhibited a strong predilection for dictionaries as expositors of ordinary meaning,[63] other justices who have relied on ordinary meaning have at times faced textualist pushback from colleagues.[64]

Stepping back, most decisions since 2000 involving interpretation of "any" have had relatively modest stakes when compared with the high-profile cases that have defined and deeply divided the Court in political and ideological terms.[65] Perhaps relatedly, the justices have played their textualist cards with a certain collegial enthusiasm. Digging into dictionary definitions, maxims of word

61. *See, e.g.*, Bostock v. Clayton Cnty. Georgia, 140 S. Ct. 1731 (2020) (Gorsuch versus Alito and Kavanaugh); Niz-Chavez v. Garland, 141 S. Ct. 1474 (2021) (Gorsuch versus Kavanaugh); Biden v. Texas, 142 S. Ct. 2528 (2022) (Roberts versus Alito); Van Buren v. United States, 142 S. Ct. 1648 (2022) (Barrett versus Thomas).

62. *See* Patel v. Garland, 142 S. Ct. 1614 (2022) (Barrett versus Gorsuch).

63. *See, e.g.*, Taniguchi v. Kan Pac. Saipan Ltd., 566 U.S. 560 (2012) (Justice Alito surveyed fourteen dictionary definitions of "interpreter"); Mac's Shell Serv., Inc. v. Shell Oil Prods., 559 U.S. 175 (2010) (Justice Alito relied on two dictionary definitions of "terminate" and also two definitions of "cancel"); *Bostock*, 140 S. Ct. at 1731 (Justice Alito in dissent included an Appendix with seven dictionary definitions of "sex" from the period on or before 1964); Sw. Airlines Co. v. Saxon, 142 S. Ct. 1783 (2022) (Justice Thomas relied on two dictionary definitions of "worker" and two definitions of "engaged"); Janus Cap. Grp. v. First Derivative Traders, 564 U.S. 135 (2011) (Justice Thomas relied on two dictionary definitions of "make": definition 59 in the 1933 Oxford English Dictionary and definition 43 in the 1934 Webster's New International Dictionary); Gross v. FBL Fin. Servs., Inc., 557 U.S. 167 (2009) (Justice Thomas relied on three dictionary definitions of "because of").

64. A recent example is *Wooden v. United States*, 142 S. Ct. 1063 (2022). Justice Kagan for the majority relied on the ordinary meaning of "occasion," citing dictionary definitions for support. *Id.* at 1069. Justice Gorsuch, concurring in the judgment, criticized Kagan's ordinary meaning analysis and invoked the Rule of Lenity as a preferred grounds for the decision. *Id.* at 1080–82. *See also* Niz-Chavez v. Garland, 141 S. Ct. 1474 (2021) (Justice Gorsuch for majority relied on the meaning of "a" in clause "a notice to appear" under 1996 immigration statute; Justice Kavanaugh in dissent criticized majority for applying what he called literal meaning rather than ordinary meaning).

65. Of more than a dozen "any" cases decided since the early 1990s and discussed in this Part, *Cir. City Stores, Inc. v. Adams*, 532 U.S. 105, 114–15 (2001) and *Massachusetts v. EPA*, 549 U.S. 497 (2007) had predictably substantial policy consequences. Most other decisions involved discrete individuals seeking to avoid criminal prosecution or imprisonment, or to pursue civil remedies of various kinds.

association and statutory structure, and substantive canons,[66] they at times almost seem to be having too much fun while the parties seek to preserve their liberty or recover a modicum of dignity.[67] Yet because the justices substantially accept a dominant methodological approach to statutory interpretation ("we are all textualists now"), they engage with one another on what has become common legal ground—in stark contrast to the partisan and ideological chasms that divide them when the Court acts as a political branch.[68]

Over this same period of two-plus decades, the justices have largely eschewed reliance on legislatively created resources such as legislative history and statutory purpose. Whether this approach is an appropriate way to manage the complexities associated with congressional reliance on certain key language—here, the word "any" in a range of settings—is worthy of deeper inquiry.

III. DEEPER DIVES

What follows below is an effort to pivot from a high-altitude perspective of "any" in the Court over time to a more ground-level analysis through pairwise comparisons of specific cases. We believe these case studies support two arguments. First, the Court's willingness to embrace a literalist approach to ordinary meaning analysis is not going to produce reliable conclusions about the scope of "any" in any given statute. Indeed, limiting the arsenal of

66. Cases in point may be the spirited contests between Ginsburg and Kagan in *Yates v. United States*, 574 U.S. 528 (2015); between Breyer and Thomas in *Small v. United States*, 544 U.S. 385 (2005); and the extended exchanges among Thomas, Kennedy, and Breyer in *Ali v. Fed. Bureau of Prisons*, 552 U.S. 214 (2008).

67. *See Yates*, 574 U.S. at 553 (Kagan, J., dissenting) (relying on Dr. Seuss's *One Fish Two Fish Red Fish Blue Fish* as textual support for broad scope of "tangible object"). *See also Bostock*, 140 S. Ct. at 1755 (Alito, J., dissenting) (describing the majority opinion as "like a pirate ship . . . sail[ing] under a textualist flag"); *Ali*, 552 U.S. at 243–44 (Breyer, J., dissenting) (relying on an imagined (?) conversation with his wife about there not being any butter in their refrigerator as support for a narrow reach of "any"). One has to wonder if Justice Breyer had recently been reading NORA EPHRON, HEARTBURN 20–21 (1996) (noting that sometimes husbands say, "'Is there any butter?' . . . We all know whose fault it is if there isn't, don't we."). Thanks to Ben Zipursky for the literary reference.

68. *See generally* John O. McGinnis, *Our Two Supreme Courts*, LAW & LIBERTY (May 6, 2015), https://lawliberty.org/our-two-supreme-courts (commentary based on Justice Kagan's 2015 speech at Northwestern Law School, in which the Justice contrasted high-profile "political" cases featuring very brief discussion at conference as everyone just states their position, with other "legal" cases in which the justices deliberate at greater length and can persuade one another).

contextual clues to only literal meaning or only intrinsic sources will likely continue to leave the jurisprudence of "any" in its current state of confusion and indeterminacy. Second, demonstrating how even a textualist Court has marshalled context—including context beyond the linguistic dimensions of the text—supports our alternative argument. The most promising avenue for the Court to pursue in deciding what a statute that uses "any" means is to admit a wide range of sources to ascertain statutory plan or purpose.

Part III.A looks carefully at two cases, decided in quick succession, that reach divergent judgments on uses of "any" in contiguous sections of the same statute. Part III.B explores two cases in which the Court overtly considered the use of the linguistic canon of *ejusdem generis*, coming to different conclusions about its application to limit a use of "any." Part III.C examines two cases applying different substantive canons as presumptions to help identify the appropriate use of "any" in a statutory scheme. Part III.D shows how purposive analysis can continue to be relevant even for a largely textualist Court that has not yet fully appreciated the inherent variety of legislative uses of the word "any."

A. Parallel Exceptions Within a Single Statute

Legislators often use "any" as part of parallel or related provisions within a single statute, in connection with shaping coverage and exclusions. This is understandable given that an underlying statutory purpose may require specifying how a broad legal principle or standard will apply, or be limited, in a range of enumerated settings. The Federal Tort Claims Act (FTCA) provides a classic example.[69]

The FTCA, enacted in 1946, abrogates the traditional doctrine of sovereign immunity by permitting private parties to sue the United States in federal court for most torts committed by persons acting on behalf of the government. Section 2680, however, contains various exceptions to this waiver, all starting with the phrase "any claim arising out of" and reflecting Congress's purpose that immunity be preserved for particular areas of activity or function by federal officials. The two exceptions of interest here are the subsection dealing with torts involving U.S. Postal Service (USPS) activities (§ 2680(b)) and the subsection addressing detention of

69. Federal Tort Claims Act, Pub. L. No. 79-601, 60 Stat. 812 (1946).

goods, merchandise, or other property by certain federal officers (§ 2680(c)). Both subsections follow the "any claim" opener with specific descriptors linked to a more general phrase of uncertain meaning. For activities involving the USPS, the FTCA preserves immunity for "any claim arising out of the loss, miscarriage, *or negligent transmission*" of postal matter.[70] For actions that involve detaining private property, the statute sustains immunity for "detention of any goods, merchandise, or other property by any officer of customs or excise *or any other law enforcement officer* . . . ."[71]

The parallels in structure distinguish these two waiver exceptions from most other provisions of § 2680, where the "any claim" is followed by a descriptor phrase that is unambiguously broad and unqualified.[72] Other lengthier FTCA waiver exceptions involve phrases that on their face have uncertain limits or ramifications,[73] but they do not follow the compact parallel structure of subsections (b) and (c).

Those two exceptions were the focus of the Roberts Court cases decided two terms apart. *Dolan v. U.S. Postal Service*[74] addressed the essence and scope of "negligent transmission" in § 2680(b); and *Ali v. Federal Bureau of Prisons*[75] analyzed the scope of "any other law enforcement officer" in § 2680(c). Our interest here is in how courts—the Supreme Court and the lower courts—construed the two contested general phrases, and how the word "any" should be analyzed in those settings.

Because the two cases concern the scope of waivers of sovereign immunity, a venerable substantive canon might be relevant to the inquiry. This canon provides that such waivers must be

70. 28 U.S.C. § 2680(b) (emphasis added).

71. 28 U.S.C. § 2680(c) (emphasis added).

72. *See, e.g.*, 28 U.S.C. § 2680(f) (no waiver for damages caused during a quarantine by the federal government), (k) (no waiver for claim arising in foreign country), (l) (no waiver for claim arising from activities of Tennessee Valley Authority), (n) (no waiver for claims arising from activities of a federal land bank or intermediate credit bank). *See also* 28 U.S.C. § 2680(e) (no waiver for claim arising out of act or omission by federal employee administering national defense provisions under title 50 of the U.S. Code), (j) (no waiver for claim arising from combatant activities during time of war), (d) (no waiver for claim where a remedy is provided under certain chapters of Title 46 of the U.S. Code relating to admiralty).

73. *See* 28 U.S.C. § 2680(a) (no waiver for claims by officials exercising due care or performing a discretionary function), (h) (no waiver for claims arising out of a range of intentional torts, with a proviso of exposure to liability for some of those torts if committed by investigative or law enforcement officers).

74. Dolan v. U.S. Postal Serv., 546 U.S. 481 (2006).

75. Ali v. Fed. Bureau of Prisons, 552 U.S. 214 (2008).

"unequivocally expressed" and are to be narrowly construed in favor of the sovereign.[76] At the same time, the FTCA is itself a broad waiver of sovereign immunity, as recognized by the Court in its practice of "narrowly constru[ing] exceptions to waivers of sovereign immunity where that was consistent with Congress' clear intent, as in the context of the 'sweeping language' of the Federal Tort Claims Act."[77] Further, when rejecting the government's sovereign immunity argument in a non-FTCA case involving textual ambiguity, the Roberts Court declared that the canon "is [simply] a tool for interpreting the law, and we have never held that it displaces the other traditional tools of statutory construction."[78] We take up the interface between "any" and substantive canons more directly in Part III.C.

1. Dolan v. United States Postal Service

Dolan involved a postal customer injured when she tripped over mail negligently left on her porch by a mail carrier. The question was whether "negligent transmission" covered such a situation as part of the exception to the waiver of immunity, or whether instead the phrase applied only to conduct directly related to "loss and miscarriage"—that is, causing mail to be destroyed or misplaced, or delivered to a wrong address. A court might decide to construe "any" expansively, so that "*any* claim arising out of . . . negligent transmission" covers all possible instances of negligent transmission by postal officials.[79] Apart from an argument based on the ordinary meaning of "any," "negligent transmission" itself

76. *See, e.g.*, United States v. Nordic Vill. Inc., 503 U.S. 30, 33–34 (1992); United States v. King, 395 U.S. 1, 4 (1969). *See generally* ANTONIN SCALIA & BRYAN A. GARNER, READING LAW 281–89 (2012) (discussing the sovereign immunity substantive canon).

77. *Nordic Vill.*, 503 U.S. at 34 (citing to several precedents). *See* Kosak v. United States, 465 U.S. 848, 855–56 (1984) (discussing FTCA's purpose of waiving sovereign immunity, with reference to § 2680(b)–(c)). *See also* SCALIA & GARNER, *supra* note 76, at 283 (describing the FTCA as having "largely eliminated" sovereign immunity with respect to tort claims).

78. Richlin Sec. Serv. Co. v. Chertoff, 553 U.S. 571, 589 (2008). The Court has been more generous reviewing how the sovereign immunity waiver should be applied once a statute is clear it has been waived. *See, e.g.*, Gomez-Perez v. Potter, 553 U.S. 474, 491 (2008). *See generally* Gregory C. Sisk, *Twilight for the Strict Construction of Waivers of Federal Sovereign Immunity*, 92 N.C. L. REV. 1245 (2014) (discussing four cases in 2012 Term where Court invoked text, context, and legislative history without relying on narrow construction favoring the sovereign).

79. *See Dolan*, 546 U.S. at 495 (Thomas, J., dissenting) (emphasis in original); Gager v. United States, 958 F. Supp. 494, 496 (D. Nev. 1997); Robinson v. United States, 849 F. Supp. 799, 801 (S.D. Ga 1994).

arguably encompasses *consequences* flowing from acts that are themselves negligent.[80] This broader understanding is also grounded in practical realities, because the normal operation of the USPS includes not simply the sorting and transfer of the mail but also the ultimate act of delivering millions of letters and packages for which Congress arguably did not mean to attach liability.[81] And even if such a reading is not fully persuasive, the postal exception could at least be deemed sufficiently ambiguous as to the scope of "any . . . negligent transmission," allowing a court to invoke the canon favoring narrow construction of sovereign immunity waivers.[82]

Alternatively, a court might take guidance from the Oregon and Delaware drafting manuals we discussed in Part I: with respect to a provision that is nonexclusive but exemplary, the reason to mention particular examples is to contextualize and shape the contours of the general words with which they are associated.[83] The Supreme Court in fact adopted this approach in *Dolan*, applying the *ejusdem generis* canon so that the meaning of "negligent transmission" covered only negligence directly related to loss or miscarriage of the mail.[84] As part of its reasoning, the Court also relied on the "meaningful variation" element of the "whole act" rule, contrasting subsection (b) on postal matters with numerous other FTCA waiver exceptions that "paint with a far broader brush."[85] In addition, the majority deflected efforts to rely on the sovereign immunity canon, arguing that unduly generous interpretations of this exception risk undermining the central purpose of the FTCA—to waive governmental immunity in sweeping language.[86] Finally, the Court addressed the practical consequences raised by the government's concern for millions of frivolous "slip-and-fall" claims related to delivery. The majority regarded the risk of such claims as analogous to risks faced by

80. *See Dolan*, 546 U.S. at 497 (Thomas, J., dissenting).

81. *See* Dolan v. U.S. Postal Serv., 377 F.3d 285, 288 (3d Cir. 2004) (citing to legislative history) [hereinafter *Dolan CA3*]; *Gager*, 958 F. Supp. at 496.

82. *See Dolan*, 546 U.S. at 497–98 (Thomas, J., dissenting); *Dolan CA3*, 377 F.3d at 287–88.

83. *See supra* Part I, text accompanying notes 27–34.

84. *See Dolan*, 546 U.S. at 486–87 (with Kennedy authoring the majority opinion). *See also* Raila v. United States, 355 F.3d 118, 120 (2d Cir. 2004) (relying on companion canon, *noscitur a sociis*, to reach same result).

85. *See Dolan*, 546 U.S. at 489–90 (citing to § 2680(i), (j), (k), (l), (m), (n)).

86. *See id.* at 491–92 (relying inter alia on *Kosak v. United States*, 465 U.S. 848 (1984)).

private businesses under state tort law, and it deemed sufficient the ordinary protections available against frivolous litigation.[87]

2. Ali v. Federal Bureau of Prisons

Two years later, in *Ali v. Federal Bureau of Prisons*,[88] the Court had to interpret the meaning of "any other law enforcement officer" under § 2680(c). In *Ali*, a federal prisoner sued the Bureau of Prisons (BOP), alleging that during his transfer between prisons, BOP officers had lost several of his personal items, some of religious and nostalgic significance. The issue was whether these officers were immune from tort liability as "any other law enforcement officers."

The Court could have adopted the same contextual analysis as in *Dolan*, concluding that congressional drafters included specific reference to customs and excise officers so as to restrict the scope of the general words "any other law enforcement officer" to officers concerned with customs and taxes.[89] The Court might have found further contextual support by invoking statutory purpose. A key reason for carving out waiver exceptions in § 2680 was to preclude tort actions where other remedies already existed.[90] And while such remedies had long existed against customs and excise officers, the same was not true for federal law enforcement officers outside the realm of customs and excise.[91]

On the other hand, certain contextual arguments favoring the government on the scope of "any" in § 2680(c) were not present

87. *See id.* at 491.

88. Ali v. Fed. Bureau of Prisons, 552 U.S. 214 (2008).

89. *See, e.g.*, *Ali*, 552 U.S. at 234–35 (Kennedy, J., dissenting) (relying on *ejusdem generis*); Andrews v. United States, 441 F.3d 220, 223–24 (4th Cir. 2006) ("Section 2680(c) presents a textbook *ejusdem generis* scenario"); Ortloff v. United States, 335 F.3d 652, 658–59 (7th Cir. 2003) (relying on *ejusdem generis* and *noscitur a sociis*). *See also* ABC v. DEF, 500 F.3d 103, 107 (2d Cir. 2007) (relying on rule against surplusage, to avoid reducing to a nullity Congress's inclusion of "officer of customs and excise").

90. *See, e.g.*, *Ali*, 552 U.S. at 246 (Breyer, J., dissenting); *Kosak*, 465 U.S. at 858; Bazuaye v. United States, 83 F.3d 482, 484–85 (D.C. Cir. 1996); *Ortloff*, 335 F.3d at 659.

91. *See Bazuaye*, 83 F.3d at 485–86 (describing longstanding availability of common law remedies against customs officers for negligently damaging detained goods, and against excise officers for improper seizure of money or property, followed by nineteenth century federal laws allowing officers sued over their excise and customs work to be indemnified, which transformed suits against these officers to suits against the government; then contrasting this history with situation of plaintiffs injured by federal officers acting in general law-enforcement capacities, who had no way to recover in a suit against the government); *Ali*, 552 U.S. at 246 (Breyer, J., dissenting) (pointing to *Bazuaye* court's analysis detailing this history).

with respect to § 2680(b) in *Dolan*. One involves a "whole act" reference to § 2680(h), which includes a definition of "law enforcement officer" as part of establishing a waiver exception for claims alleging libel, slander, and related torts by such officers.[92] This definition could be construed *in pari materia* with the reference to "other law enforcement officers" in § 2680(c).[93] A separate, broader contextual clue along the lines of what is sometimes called the "whole code" rule[94] is that BOP officials are considered "law enforcement officers" under a range of other federal statutes.[95] Such an argument invokes an understanding that law enforcement officers should be deemed covered under § 2680(c) without the need for further definition because the rest of the U.S. Code confirms that coverage.

More support is available from extrinsic sources that the modern Court rarely likes to consult. A central document in the legislative history of § 2680(c), previously relied on by the Supreme Court, arguably identifies a broader purpose behind that section because of its "special reference to the detention of imported goods in appraisers' warehouses or customs houses, *as well as seizures by law enforcement officials, internal revenue officers, and the like*."[96] Thus, given the contested arguments based on context, structure, and history, a court might have a sounder basis for contending that there is truly sufficient ambiguity as to the scope of

92. *See* 28 U.S.C. § 2680(h) (defining "investigative or law enforcement officer" for purposes of that subsection as "any officer of the United States who is empowered by law to execute searches, to seize evidence, or to make arrests for violations of Federal law").

93. *See* Chapa v. U.S. Dep't of Just., 339 F.3d 388, 390 (5th Cir. 2003); Bramwell v. U.S. Bureau of Prisons, 348 F.3d 804, 806–07 (9th Cir. 2003).

94. *See* Anita S. Krishnakumar, *Cracking the Whole Code Rule*, 96 N.Y.U.L. REV. 76, 76 (2021); William W. Buzbee, *The One-Congress Fiction in Statutory Interpretation*, 149 UNIV. PA. L. REV. 171, 173 (2000).

95. *See Chapa*, 339 F.3d at 390 (noting that Congress has identified BOP employees as "law enforcement officers" when determining eligibility for Civil Service premium pay and retirement benefits, and that someone who fatally injures a BOP employee while he is engaged in official duties may be charged with the felony of killing a "law enforcement officer"); *Bramwell*, 348 F.3d at 807 (same). *See also* Ysasi v. Rivkind, 856 F.2d 1520, 1525 (Fed. Cir. 1988) (holding that border patrol division employees of the Immigration and Naturalization Service who detained appellant's truck were "law enforcement officers" under § 2680(c)).

96. Kosak v. United States, 465 U.S. 848, 856 (1984) (emphasis added) (quoting from Alexander Holtzhoff's Report on Proposed Federal Tort Claims Bill). This legislative history was quoted with approval in *Ysasi*, 856 F.2d at 1525.

§ 2680(c) to consider application of the sovereign immunity canon, in the final analysis.[97]

Yet the Supreme Court majority in *Ali* invoked none of the structural or purposive arguments relied on by other courts to support the government's position. Instead, Justice Thomas for the Court focused on the use of "any" accompanying "other law enforcement officers," emphasizing the expansive dictionary meaning of "any" and citing to prior Supreme Court precedents that had construed "any" in open-ended terms.[98] The majority took time to deflect petitioner's reliance on *ejusdem generis*, *noscitur a sociis*, and the rule against superfluities;[99] and it also asserted support for its position based on language added to § 2680(c) in 2000.[100] But in the end, Justice Thomas relied on his plain meaning analysis of "any," concluding that "we are unpersuaded by petitioner's attempt to create ambiguity where the statute's text and structure suggest none[,]" and that "'any other law enforcement officer' [should be] read to mean what it literally says."[101] Preferring literal meaning to ordinary meaning, the Court revealed a confidence that it needs limited sources to derive the meaning of "any."

These two FTCA decisions illustrate the importance of linguistic, structural, and pragmatic context in addressing so-called textual disputes that implicate "any." In *Dolan*, disputes about plain meaning were eclipsed by these larger contextual considerations. The majority in effect tracked the drafting guidance set forth in certain state manuals regarding nonexclusive exemplary enumerations, supported by responses to the practical argument advanced by the government. Justice Thomas, in solitary dissent,

97. *See Chapa*, 339 F.3d at 391.

98. *See* Ali v. Fed. Bureau of Prisons, 552 U.S. 214, 218–20 (2008).

99. *See id.* at 223–26.

100. *See id.* at 221–23. The Civil Asset Forfeiture Reform Act of 2000 (CAFRA) added four subsections to § 2680(c), extending the waiver of sovereign immunity to settings where the goods injured or lost were detained as part of a property seizure "for the purpose of forfeiture *under any provision of Federal law*." 28 U.S.C. § 2680(c)(1) (emphasis added). According to the majority, the 2000 CAFRA amendment confirmed the earlier scope of "law enforcement officer" as covering all officers who implement forfeiture laws, *Ali*, 552 U.S. at 222. Justice Kennedy in dissent contended that the amendment was fully consistent with the narrower scope of the waiver exception—because customs and excise officers effect forfeitures under all laws, not just customs and excise laws, the amendment simply extended their immunity to forfeiture actions taken pursuant to laws outside the customs/excise ambit.

101. *Id.* at 227–28.

relied on the asserted breadth of "any claim" in the postal exception,[102] but he too contended that in practical terms negligence among postal carriers extends well beyond the two enumerated examples.[103]

The Court in *Ali* reflects a more rigidly textualist approach to "any." Especially striking is the majority's conclusion that there is no ambiguity at all about an expansive reading, based principally on a dictionary definition and cites to prior Court decisions. Yet string-cite appeals to precedent on "any" have limited value, given that, as we showed earlier, there are innumerable Supreme Court cases on both sides of the line.[104] In this instance, contextual arguments against an expansive reading of "any" included reference to canons on which congressional as well as state legislative drafters often rely,[105] as well as appeals to the purpose of the customs and excise exception itself. And even defenders of an expansive reading made a decent case in the circuit courts for contextual coverage of all law enforcement officers, or at least ones in the BOP.[106] While there is pushback available against these context-driven contentions for an expansive interpretation,[107] it is at that level of context—be it canons of word association, canons of structural integrity, canons about sovereign immunity, congressional purpose, or, better, all of the above—that judicial battles should be waged. In this instance, it is disappointing that the Supreme Court majority embraced literal meaning as dispositive, while essentially eschewing more meaningful context from intrinsic and extrinsic sources.

---

102. *See* Dolan v. U.S. Postal Serv., 546 U.S. 481, 495 (2006).

103. While Thomas argued for sufficient ambiguity to invoke the sovereign immunity canon, the majority emphasized that the canon is "unhelpful" in the FTCA setting. *Id.* at 491–92 (citing Kosak v. United States, 465 U.S. 848 (1984)).

104. *See supra* Parts II.A & II.B.

105. *See supra* Part I for more on state drafting manuals and the canons—and for more information about drafters' perspectives on canons in Congress, see generally Gluck & Bressman, *supra* note 35.

106. *See supra* notes 92–97, and accompanying text.

107. For example, regarding reliance on 28 U.S.C. § 2680(h) as *in pari materia* with its definition of "law enforcement officer," the definition is deemed "for the purpose of this subsection" hence perhaps not expandable to 28 U.S.C. § 2680(c). *See* Andrews v. United States, 441 F.3d 220, 226 (4th Cir. 2006). With respect to other elements of the U.S. Code identifying BOP employees as "law enforcement officers," no federal statute mentions the BOP in the context of property detention whereas out of nine federal statutes besides § 2680(c) referring to detention of goods, the majority are specific to customs and excise. *See* Ali v. Fed. Bureau of Prisons, 552 U.S. 214, 236–37 (2008). (Kennedy, J., dissenting).

B. Tension Between Internal and External Contexts

As our discussion of *Dolan* and *Ali* makes clear, debates about "any" often go hand-in-hand with arguments about the application of the linguistic canon of *ejusdem generis* (and sometimes its sister canon, *noscitur a sociis*). That is sensible, as the canon is traditionally applied so that general words in statutes (which often come with an "any" attached) are limited to some property holding together the enumerations of specific items that follow or precede the general category. The two case studies below take different approaches to the relevance of *ejusdem generis* in limiting the word "any." Watching the two cases struggle with the issue side by side enables us to show the kinds of sources the Court has found relevant to figure out when "any" can be limited by the textual canon and when it cannot. Although the approach we describe in the first case feels like a relic from a time when the Court viewed legislative history as much more relevant to statutory interpretation, it is nevertheless instructive because the problem of knowing how to interpret "any" isn't going away anytime soon. The Court needs a more systematic interpretive approach, with more tools in its arsenal than looking up "any" in the dictionary, or referencing ordinary meaning when actually implementing literal meaning with the use of italics and string cites.

1. Harrison v. PPG Industries, Inc.[108]

This case was about a jurisdictional provision of the Clean Air Act Amendments of 1977 that seemed to expand direct review in the Courts of Appeals for Administrator decisions at the Environmental Protection Agency (EPA). Section 307(b)(1) of the Act provides that a petitioner may have the Court of Appeals rather than a federal district court review "any other final action of the Administrator under [the Act] . . . which is locally or regionally applicable."[109] This provision follows a more specific enumeration of particular types of petitions, none of which covered PPG Industries' (PPG) exact circumstance. The Regional Administrator of the EPA had notified PPG by letter that a part of a PPG facility would be subject to regulation under the Act as a

108. Harrison v. PPG Indus., Inc., 446 U.S. 578 (1980).

109. 42 U.S.C. § 7607(b)(1).

statutory "new source."[110] Thus, the question arose for PPG whether it could seek review of that informal letter determination in the Court of Appeals rather than the district court under the "any other final action" section of the jurisdictional grant to the appellate court.

The Court held, italicizing "any" thrice,[111] that the petitioners were correct in their "literal" rendering to have the jurisdictional grant reach any Administrator decision, so long as it was a "final" decision that "is locally or regionally applicable."[112] The respondents' effort to get a limiting reading of "any" by looking to what might hold the prior enumerations in the statute together to constrain the domain of "any"—in an *ejusdem generis* application—was, the Court held, "misplaced."[113]

The law of *ejusdem generis* in *Harrison* is in some ways familiar to modern eyes but in other ways instructive about a different way of doing business from how the modern Court has much more summarily dealt with "any" problems. Familiarly, the Court looked to see if there really was some relevant property that tied together the prior enumerations in § 307(b)(1). The respondents had urged the Court to limit "any" because the vast majority of the enumerated kinds of Administrator decisions that were subject to direct review in the Court of Appeals were ones that would have had "a contemporaneously compiled administrative record . . . based on administrative proceedings reflecting at least notice and opportunity for hearing."[114] And the PPG letter didn't have a robust administrative record associated with it, implicating the competence of the appeals court to review the decision; it would make more sense, the respondents argued, to have these kinds of informal adjudications reviewed first in a trial court, which has more capabilities to develop a factual record. But the Court found a fatal flaw with the argument: "at least one of the specifically enumerated provisions in § 307(b)(1) . . . does not require the Administrator to act only after notice and opportunity for a hearing."[115] Accordingly, this property could not be the basis to limit "any" along the lines

110. 42 U.S.C. § 7411.

111. *Harrison*, 446 U.S. at 586, 587, 589.

112. *Id.* at 587.

113. *Id.* at 587–88.

114. *Id.* at 587.

115. *Id.* at 588.

of *ejusdem generis*. That is a common enough way to do *ejusdem generis* business.

But less common for today's Court was the second "more fundamental" reason the Court rejected the application of *ejusdem generis* in *Harrison*. The Court found there was no uncertainty created by the legislative history about the reach of "any" that would explain or justify limiting its application through a textual canon.[116] The Court reviewed the legislative history extensively—how Congress added the language in the 1977 amendment process as well as discussions in the House Committee on Interstate and Foreign Commerce about the statute's coverage and its relationship to proposals and recommendations made by the Administrative Conference.[117]

The Court also considered and rejected "the theory of the dog that did not bark." Respondents had urged that the absence of extensive and explicit legislative history about such a radical change to the jurisdiction of the Court of Appeals and a stripping of jurisdiction of the district courts during deliberations of the 1977 amendments ought to be treated as evidence that "any" should be read in a limited way. The Court conceded that the legislative history it studied and summarized was not smoking gun evidence about the matter in controversy. Still, the Court ultimately thought the expansion of jurisdiction "would not appear so large as ineluctably to have provoked comment in Congress" and that the text was clear enough to require legislative history to make it more ambiguous—rather than using absence of legislative history in the way the respondents preferred.[118] Justice Blackmun's concurring opinion underscored the point: although he found "it difficult to believe that Congress would undertake such a massive expansion in the number of Agency actions directly reviewable by the courts of appeals," he "[n]onetheless" agreed "with the Court that the dearth of evidence to the contrary makes its broad interpretation" of "any" "inescapable."[119]

To sum up *Harrison*, then, we can say this: The Court is willing to use *ejusdem generis* to constrain a reading of "any"—but only if

116. *Id.* at 588–89.

117. *Id.* at 589–91 (citing H.R. 6161, 95th Cong. (1977) (enacted); H.R. REP. NO. 95-294, at 323–24 (1977)).

118. *Id.* at 591–92.

119. *Id.* at 595 (Blackmun, J., concurring).

(1) some clear property applies to all of the specific enumerations in the provisions, or (2) some uncertainty about the coverage of "any" presents itself either in the text or can be implied from the legislative history. Absence of legislative history might be relevant if a drastic change would have "ineluctably . . . provoked comment in Congress." But the size of the jurisdictional shift encompassed by "any other agency action" was not of that magnitude—thus "dogs not barking" was ultimately irrelevant. Admittedly, the modern Court has a much more allergic reaction to the use of legislative history (and its absence!). Still, we need to bring more order to the problem of the statutory "any"—and *Harrison* points one way toward coherence, drawing from intrinsic and extrinsic sources. Even the primary dissent in *Harrison*—Justice Rehnquist's—took legislative history to be essential to understanding "any," although he disagreed about which parts of the legislative history could point the way to the right answer and how much to use the "dog didn't bark" theory, rejected by the majority.[120]

2. Circuit City Stores, Inc. v. Adams[121]

A case like *Harrison* is hard to imagine in our current textualist Court, one that tends to eschew careful interrogations of legislative history. *Circuit City* then presents a more contemporary example of how "any" and *ejusdem generis* interface—though one that offers a cautionary tale for those unwilling to try to understand the legislative record.

This case focused on the Federal Arbitration Act's[122] (FAA) first section, which excludes from the FAA's coverage "contracts of employment of seamen, railroad employees, or any other class of workers engaged in foreign or interstate commerce."[123] The Ninth Circuit had held—in conflict with all the other circuits to have considered the issue—that an arbitration agreement in an

120. *Id.* at 595–602 (Rehnquist, J., dissenting). The Rehnquist dissent in *Harrison* was drawn upon regularly by Justice Stevens when he deployed the "dog-didn't-bark" canon of statutory interpretation. *See* Chisom v. Roemer, 501 U.S. 380, 396 n.23 (1991); Griffin v. Oceanic Contractors, Inc., 458 U.S. 564, 589 n.20 (1982) (Stevens, J., dissenting). Ironically perhaps, Stevens did not join the Rehnquist dissent in *Harrison* and wrote his own, one that was much less informed by a deep dive into the legislative history. *Harrison*, 446 U.S. at 602–07 (Stevens, J., dissenting).

121. Cir. City Stores, Inc. v. Adams, 532 U.S. 105 (2001).

122. 9 U.S.C. § 1 (1925).

123. *Id.*

employment contract was not subject to the FAA's preemption of state law because a "contract of employment" comes within the exclusion language of "any other class of workers engaged in . . . interstate commerce."[124] The Ninth Circuit explicitly rejected the application of *ejusdem generis*—an important part of some other circuit court engagements with the question[125]—arguing instead that (1) the § 1 exclusion had to be read in light of § 2's coverage section; and, more importantly, (2) that the textual canon could not be used to defeat the clear legislative history and the legislative purpose of the FAA.[126]

The Supreme Court's conservative, then-five-justice majority in 2001 signed onto Justice Kennedy's *ejusdem generis* argument reversing the Ninth Circuit.[127] By 2001, Justices Kennedy, Thomas, O'Connor, Scalia, and Rehnquist were seemingly willing to consider "statutory context . . . in a manner consistent with the FAA's purpose."[128] But their engagement focused rather narrowly on "the text" in isolation from the "legislative history of the exclusion provision"[129]—from which they might have understood more context and more nuanced purpose that was Congress's rather than the Supreme Court's own gloss thereupon.[130] And although by the opinion's end the Court offered a superficial engagement with what it called the "sparse" "legislative record on the § 1 exemption,"[131] the majority also very clearly found the *ejusdem generis* argument to be an "insurmountable textual obstacle"[132] that immunized it even from seriously "assess[ing] the legislative history."[133]

124. *See* Craft v. Campbell Soup Co., 177 F.3d 1083 (9th Cir. 1999). *See also* Cir. City Stores, Inc. v. Adams, 194 F.3d 1070 (9th Cir. 1999) (applying *Craft* to the contract between Circuit City and Adams).

125. *See, e.g.*, Tenney Eng'g v. United Elec. Radio & Mach. Workers of Am., 207 F.2d 450, 452 (3d Cir. 1953) (applying *ejusdem generis* to Section 1 of the FAA); Cole v. Burns Int'l Sec. Servs., 105 F.3d 1465, 1471 (D.C. Cir. 1997) (citing Harrison v. PPG Indus., Inc., 446 U.S. 578, 588 (1980)) (same).

126. *Craft*, 177 F.3d at 1092.

127. Cir. City Stores, Inc. v. Adams, 532 U.S. 105, 114 (2001).

128. *Id.* at 118.

129. *Id.* at 118–19.

130. As Justice Stevens argues in dissent, "the Court is standing on its own shoulders" when it develops its purposive readings of the FAA. *Id.* at 132 (Stevens, J., dissenting).

131. *Id.* at 119–21 (majority opinion).

132. *Id.* at 114.

133. *Id.* at 119.

The Court essentially refused to resort to legislative history "to cloud a statutory text that is clear."[134] But as Justice Souter urged in his dissent for four justices, "the Court has repeatedly explained that the [*ejusdem generis*] canon is triggered only by uncertain statutory text . . . and that it can be overcome by, *inter alia*, contrary legislative history . . . . The Court [in the majority opinion] turns this practice upside down, using *ejusdem generis* to establish that the text is so clear that legislative history is irrelevant."[135]

Justice Stevens's dissent (for three justices) walked through the legislative history for the Court. He found this history favored the Ninth Circuit's reading because of how central *commercial* arbitration was in the bill that led to the FAA.[136] Moreover, accusing the Court of "[p]laying ostrich,"[137] Stevens described how the amendment that became § 1 was introduced by the bill's principal supporters in order to remove opposition from organized labor over the possibility that the bill would cover labor disputes or workers' contracts.[138]

Justice Souter's dissent also added some legislative history to the mix to support the Ninth Circuit,[139] and offered a counter-canon of *ex abundanti cautela*: Congress sometimes lists specifics not to limit general terms but rather to make sure the enumerated items get covered with belts-and-suspenders.[140] The Court, however, preferred to use *ejusdem generis* to assert textual clarity that couldn't be defeated by substantial evidence of congressional intent. That isn't where the Court was in the days of *Harrison* but may be where a more textually oriented Court will find itself today. That is a shame because, for all the lip service to "statutory context," a Court playing ostrich about the congressional record and scheme—refusing "to look beyond the raw statutory text"[141]—will find it

134. *Id.* (quoting Ratzlaf v. United States, 510 U.S. 135, 147–48 (1994)).

135. *Id.* at 138 n.2 (Souter, J., dissenting) (citations omitted).

136. *Id.* at 125–26 (Stevens, J., dissenting).

137. *Id.* at 128.

138. *Id.* at 126–28.

139. *Id.* at 138–39 n.3 (Souter, J., dissenting) (citing *Sales and Contracts to Sell in Interstate and Foreign Commerce, and Federal Commercial Arbitration: Hearing on S. 4213 and S. 4214 Before the Subcomm. of the S. Comm. on the Judiciary*, 67th Cong. 9 (1923) (statement of Senator Walsh).

140. *See generally* Ethan J. Leib & James J. Brudney, *The Belt-and-Suspenders Canon*, 105 IOWA L. REV. 735 (2020) (citing numerous examples where Congress and state legislatures employ redundant drafting practices to reinforce the meaning of a term or provision, thereby seeking to ensure that implementing agencies and courts understand the legislative message).

141. *Cir. City*, 532 U.S. at 132 (Stevens, J., dissenting).

very challenging to figure out what "any" is doing in statutes by using only textual resources internal to the U.S. code.

C. The Impact of Policy Presumptions

Unlike the textually related canons addressed in sections A and B, which are purportedly derived from understandings about "normal uses of language by educated speakers,"[142] substantive canons are generally rooted in policy judgments based on the structure or expected scope of government within our federal constitutional system. Two such policy judgments invoked by the Court with some frequency in its statutory decisions are the presumption against federal legislation applying to extraterritorial matters[143] and the presumption against federal law preempting state authority.[144] These presumptions, to be sure, can be overcome by sufficiently clear statutory text. The two cases that follow, from the late Rehnquist Court, invoke aspects of these presumptions and construe "any" in the context of deciding whether Congress has in fact been sufficiently clear to overcome them. In each case, the Court's answer is no. The more general point, however, is that part of figuring out the function "any" serves in a given statute is to understand its embeddedness within an interpretive regime that includes some normative and policy presumptions the courts regularly consider and acknowledge.

1. Small v. United States[145]

The issue in *Small* arose under 18 U.S.C. § 922(g)(1), which prohibits any person from possessing a firearm if "convicted in any court" of a crime punishable by imprisonment for more than one year. Small was convicted in a Japanese court of trying to smuggle firearms and ammunition into that country and sentenced there to five years' imprisonment. Upon returning to the United States, he was subsequently indicted for possessing a firearm in violation of § 922. He moved to dismiss the indictment, contending that

142. SCALIA & GARNER, *supra* note 76, at 243.

143. *See, e.g.*, EEOC v. Arabian Am. Oil Co., 499 U.S. 244, 249–51 (1991); Foley Bros. v. Filardo, 336 U.S. 281, 285 (1949).

144. *See, e.g.*, Gregory v. Ashcroft, 501 U.S. 452, 459–60 (1991); United States v. Bass, 404 U.S. 336, 349 (1971).

145. Small v. United States, 544 U.S. 385 (2005).

Congress's reference to "any court" should be read to exclude convictions entered by a foreign court.

Justice Breyer for the majority began by invoking context, insisting that the word "any" considered alone could not answer the question whether foreign court convictions fall within the scope of the statute. In support of this contextual approach, Breyer quoted from Chief Justice Marshall's hoary pronouncement that "general words, such as the word 'any,' must be limited in their application 'to those objects to which the legislature intended to apply them,"[146] and added the by-now familiar string cite supporting a contextual approach to "any."[147]

For the majority, the initial policy context was the "commonsense notion that Congress generally legislates with domestic concerns in mind," and its statutes therefore are ordinarily meant to have domestic, not extraterritorial, application.[148] Breyer acknowledged that the presumption against extraterritoriality did not map neatly onto this case because the statute was not being applied outside the territorial United States. But he found a similar assumption to be appropriate when considering the scope of the phrase "convicted in any court."

It may seem counterintuitive to apply this assumption to conviction for gun possession itself, especially given that the district court had exhaustively examined the Japanese trial record and determined that Small's conviction comported with standards of fundamental fairness.[149] But the Court focused on the scope of "any" if applied to foreign convictions *as a class* rather than the factual details of this individual case. In the broader context, it concluded that for crimes punishable by more than one year's imprisonment, foreign laws might well encompass conduct that domestic law would economically encourage, or constitutionally protect, or simply penalize with less than a year in jail.[150]

Although the array of important distinctions between foreign and domestic criminal offenses supported the policy assumption

146. *Id.* at 388 (quoting United States v. Palmer, 16 U.S. (3 Wheat.) 610, 631 (1818) (internal quotation marks omitted).

147. *See supra* note 7 (highlighting that *Small* cited four prior Court decisions on "any").

148. *Small*, 544 U.S. at 388–89 (quoting Smith v. United States, 507 U.S. 197, 204 n.5 (1993)).

149. *See* United States v. Small, 183 F. Supp. 2d 755, 763–69 (W.D. Pa. 2002).

150. *See Small*, 544 U.S. at 389–90 (describing punishment for certain profitmaking conduct under Russian or Cuban law, for expressive propaganda activity under Cuban law, and for vandalism under Singapore law).

about the ordinary reach of domestic statutes, the Court emphasized that this was a presumption rather than a clear statement rule. Continuing with its examination of context, the Court stated that it was "ready to revise this assumption should statutory language, . . . history, or purpose show the contrary."[151] But the majority found no such convincing evidence to exist.

On text, the majority pointed to a series of contiguous statutory provisions that address exceptions, applications, or enhancements of the penalty for unlawful gun possession.[152] Because all of these refer only to crimes under federal or state law, the inference was that Congress simply did not consider whether the phrase "convicted in any court" applies to foreign convictions.[153] Justice Thomas in dissent invoked the expansive dictionary definition of "any" and cited to prior decisions taking a similarly literalist approach.[154]

On legislative history, the Court acknowledged that the Senate bill, rejected in conference, contained language that would have restricted predicate offenses to domestic crimes, separately identifying "Federal" crimes and "State" felonies on that score.[155] From the majority's standpoint, the Conference Committee's preference for the "convicted in any court" language reflected *not* a desire to incorporate foreign law but rather interest in a simpler approach that avoided problems should states adopt differing definitions of the term "felony."[156] Thus legislative history's silence

151. *Id.* at 391.

152. *See id.* at 391–92 (provisions allowing gun possession for certain business regulatory crimes, specifying that predicate crimes include misdemeanor crime of domestic violence, and enhancing penalties for three predicate drug offense convictions).

153. The majority explained how construing "convicted in any court" to include foreign convictions would result in various anomalies; for instance, someone convicted of a Canadian antitrust offense could not possess a gun while someone convicted under federal or state law could do so; a person convicted of three serious drug offenses under federal or state law would suffer enhanced punishment but not if convicted three times under foreign law. *Id.* at 391–92.

154. *Id.* at 396–97 (Thomas, J., dissenting). Thomas also relied on linguistic context, asserting (per *expressio unius*) that unlike other sections of the firearms-control statute expressly referencing federal or state law, § 922(g)(1)'s simple reference to "any law" signified the absence of geographic limits as to scope. *Id.* at 397–98. Justices Scalia and Kennedy joined the Thomas dissent.

155. *See id.* at 393 (majority opinion).

156. *See id.*

on foreign law was simply a neutral factor, confirming that Congress did not even consider the issue.[157]

Finally, on purpose, the majority conceded that one aspect of Congress's purpose was keeping guns out of the hands of those who had demonstrated threatening capacities, and that conviction for a serious crime abroad might well reflect such a level of threat or dangerousness.[158] The majority's rebuttal was to point to empirical evidence that foreign convictions had only been a predicate for prosecution under § 922(g)(1) about a dozen times in the nearly four decades since 1968.[159] This may be the majority's least persuasive argument: even a dozen instances where courts found foreign conviction to be a suitable predicate for firearms-possession prosecution suggests that convictions in a foreign court may be recognized as appropriate under this statutory scheme.

But even with this less persuasive element of the Court's analysis, the larger lesson from *Small* is that the context for "any" can include—and indeed be framed by—substantive policy presumptions as well as language and structural canons. In response to this policy presumption, the majority finds it essential to consider a range of interpretive factors besides literal textual meaning. Indeed, apart from Justice Thomas's dictionary-driven approach, he too seems prepared to accept this possibility, reinforcing his spare textualist position with arguments based on structural canons and legislative history.

2. Nixon v. Missouri Municipal League[160]

A section of the Telecommunications Act of 1996[161] authorizes the preemption of state and local laws if they "prohibit[] the ability of any entity" to provide telecommunications services. In 1997, Missouri went ahead and passed a law prohibiting all political subdivisions of the state from offering telecommunications

157. *See id.* For Justice Thomas, hardly known to embrace the relevance of legislative history, the Conference Committee's deletion of reference to "Federal" and "State" was probative: it was, for him, "strong confirmation of the fact that 'any' means not 'any Federal or State,' but simply 'any.'" *Id.* at 406–07 (Thomas, J., dissenting).

158. *See id.* at 393–94.

159. *See id.* at 394.

160. Nixon v. Mo. Mun. League, 541 U.S. 125 (2004).

161. 47 U.S.C. § 253.

services.[162] Municipalities in the state then asked the FCC to declare Missouri's law preempted under the statute. The FCC declined to preempt[163] and the Court had to determine in this case whether the relevant class of entities in the "any entity" locution in the law includes a state's own subdivisions, ultimately affecting a state's right to restrict its own subdivisions from delivering telecommunications services. The Eighth Circuit had reversed the agency because it found that the plain-vanilla language of "entity" modified with "any" was sufficiently clear to countermand the principle the FCC invoked from *Gregory v. Ashcroft*,[164] "that Congress needs to be clear before it constrains traditional state authority to order its [own] government."[165] But the Court ultimately reinforced the *Gregory v. Ashcroft* presumption and essentially found that "any" had to be read in light of the substantive canon and not as a quick opt-out for it.

Indeed, the Court from the outset sought to be clear that "'any' can and does mean different things depending upon the setting."[166] It was underwhelmed with the Eighth Circuit's cramped "analysis on the words 'any entity,' left undefined by the statute, with [too] much weight being placed on the modifier 'any.'"[167] Highlighting that "any entity" can mean public or private entities and that "any" can be "expansive" or "narrow," the Court reached for a "broader frame of reference" than just the proverbial "writing on the page."[168]

That broader frame was "a working assumption that federal legislation threatening to trench on the States' arrangements for conducting their own governments should be treated with great skepticism, and read in a way that preserves a State's chosen disposition of its own power."[169] Citing to *Gregory*, the Court concluded that "the ability of any entity" language in the statute was not "forthright enough" nor was it limited to only one meaning.[170] Without an "unmistakably clear statement," the

162. MO. REV. STAT. § 392.410(7) (2008) (originally enacted as L. 1997, H.B. No. 620 § A).

163. *In re* Mo. Mun. League, 16 FCC Rcd. 1157 (2001).

164. Gregory v. Ashcroft, 501 U.S. 452 (1991).

165. *Nixon*, 541 U.S. at 130 (citing 16 FCC Rcd. at 1169; Mo. Mun. League v. FCC, 299 F.3d 949 (8th Cir. 2002)).

166. *Id.* at 132.

167. *Id.*

168. *Id.* at 132–33.

169. *Id.* at 140.

170. *Id.* at 141.

Court read "any" in light of the background of traditional legal meanings pursued through a substantive canon.

Lest anyone think this mode of reading "any" is only for liberals, the case's core conclusion here was amplified in a concurrence by Justices Scalia and Thomas. They wrote separately to distance themselves from the Court's reliance in part on consequentialist analysis and explained that the substantive canon sufficiently controlled the case—and that the "any" language could not be a source of required preemption from the FCC.[171] It was Justice Stevens who dissented on grounds close to a "plain meaning" reading underwritten by some legislative history and purpose.[172]

What these two case studies illustrate is that the scope of "any" can often come from background normative priorities as evidenced by hoary substantive canons. Although we have drawn on relatively recent cases using presumptions disfavoring extraterritorial applications and interference in a state's internal dispositions of power, it is not hard to see how other canons (such as the rule of lenity[173] or the rule against interpreting statutes to be retroactive[174]) might come to inform the reach of "any" in a statutory context. The substantive canon is not in these renderings wholly extrinsic to the scope of "any" but instead helps supply context for its usage.

D. The Centrality of Purpose to the Hermeneutics of "Any"

We have thus far examined different kinds of context that can usefully guide courts about the function "any" serves in a statutory scheme; "any" is neither self-defining nor likely to be illuminated by the dictionary. Such context can come from linguistic canons like *ejusdem generis*. And it can come from other conventionally intrinsic textual sources like the "whole act" or the "whole code." It can also be supplied by a substantive canon, as we just showed. For judges willing to consider other "extrinsic" sources of context, a particular

171. *Id.* at 141 (Scalia, J., concurring).

172. *Id.* at 142–48 (Stevens, J., dissenting).

173. *See* McBoyle v. United States, 283 U.S. 25 (1931) (holding an airplane not to be "any other self-propelled vehicle not designed for running on rails" because the National Motor Vehicle Theft Act lacked "fair warning" that it would apply to theft of airplanes).

174. *See generally* Claridge Apartments Co. v. Comm'r of Internal Revenue, 323 U.S. 141, 142–43 n.1, 164–65 (1944) (holding that the statutory language of "any" was not clear enough to overcome the presumption against retroactivity).

use of "any" in a given statute can be informed by the records of deliberation in the legislature, offering guidance as to whether an expansive or limiting usage is more appropriate. But an extrinsic source of context for "any" that is somewhat more acceptable to textualists is statutory purpose. That purpose is typically gleaned from evaluating the record and background information in the lead-up to introduction and enactment, in order to establish what the statute sought to accomplish. In the two cases we discuss below, one authored by Justice Scalia and one authored by Justice Thurgood Marshall, the Court invokes purpose to figure out how "any" operates in a statutory provision. Although we don't offer these examples to sell the results reached by the Court, the two cases highlight the more general use of purpose in adjudicating the scope of "any" for judges with quite disparate methodological priorities.

1. AT&T Mobility LLC v. Concepcion[175]

Like *Circuit City*, this case also involved the proper interpretation of a part of the FAA.[176] At issue was § 2's command to render arbitration agreements "valid, irrevocable, and enforceable, save upon such grounds as exist at law or in equity for the revocation of *any* contract . . . ."[177] Because California courts had found unconscionable and therefore unenforceable class action waivers in consumer form contracts under certain conditions,[178] the lower courts applying California contract law found that the class waiver in the arbitration agreement between the Concepcions and AT&T Mobility was also unconscionable and unenforceable. After all, the public policy proscription on class waivers would have been a ground for a finding of unconscionability for *any* consumer form contract, irrespective of whether the agreement contained an arbitration clause.

Nevertheless, the Supreme Court, in an opinion by Justice Scalia, deemed the generally-applicable unconscionability law of California—that would have disallowed a class action waiver to apply to any contract whether in a litigation or an arbitration

175. AT&T Mobility LLC v. Concepcion, 563 U.S. 333 (2011).

176. 9 U.S.C. § 2.

177. *Id.* (emphasis added).

178. Discover Bank v. Superior Ct., 36 Cal. 4th 148 (2005).

context—to be preempted by the FAA. Although Justice Breyer in opening his dissent for Justices Ginsburg, Sotomayor, and Kagan used the classic strategy of italicizing "any" in the statute to try to get the point across (perhaps forgetting his nuanced view of "any" in *Ali* just three years earlier),[179] the majority opinion focused on certain extrinsic evidence to decipher the language of "any" in the FAA.

Indeed, Justice Scalia invoked legislative *purpose* five times in the majority opinion to find that the FAA preempts California contract law: the majority argued that the "overarching purpose of the FAA . . . is to ensure the enforcement of arbitration agreements according to their terms so as to facilitate streamlined proceedings."[180] Although claiming this purpose could be found in "the text of §§ 2, 3, and 4,"[181] nothing in the FAA clearly announces a purpose focused on "streamlined proceedings" or takes a position about class actions. The Court pivoted quickly from the text of the FAA to the Court's own longstanding gloss thereupon about its "principal purpose," as if recognizing it would be better to cite itself than any specific intrinsic textual source on the centrality of individual rather than class arbitration or streamlined proceedings more generally.[182] And without citation, the Court also argued that states were free to "requir[e] class-action-waiver provisions in adhesive agreements to be highlighted" so long as "[s]uch steps" don't "frustrate" the FAA's "purpose."[183]

From one perspective, there is nothing remarkable about this statement: standard "conflict preemption" analysis virtually requires investigation into legislative purpose to make sure state laws aren't serving to undermine a federal statutory scheme.[184] From another perspective, however, the willingness of the Court's conservative majority to allow congressional purposes to inform

179. *AT&T*, 563 U.S. at 357 (Breyer, J., dissenting); *see also id.* at 359 (italicizing "any" a second time). See our discussion of *Ali supra* Part III.A.

180. *AT&T*, 563 U.S. at 344.

181. *Id.*

182. *Id.* (citing Volt Info. Scis., Inc. v. Bd. of Trs. of Leland Stanford Junior Univ., 489 U.S. 468, 478 (1989)).

183. *Id.* at 347 n.6.

184. *See, e.g.*, Geier v. Am. Honda Motor Co., 529 U.S. 861 (2000) (focusing on the centrality of purpose in conflict preemption analysis). The opinion was authored by Justice Breyer joined by Justices Scalia, Kennedy, Rehnquist, and O'Connor. The dissent—by Justice Stevens joined by Justices Thomas, Ginsburg, and Souter—didn't dispute that purpose was significant for the analysis, but still relied in part on "any" in the statutory savings clause to argue against preemption in that case. *Id.* at 898 (italicizing "any," as is traditional).

its reading of "any" in statutory language underscores that the jurisprudence of "any" even for largely textualist judges can invoke purposes to help navigate the word's routine variability.

But it wasn't only Justice Breyer who forgot his "any" jurisprudence from three terms previous. Any-means-any Justice Thomas wrote a concurrence in *AT&T* that similarly departed from his views in *Ali*. Thomas simply thought "[i]t would be absurd to suggest that § 2 requires only that a defense apply to 'any contract.'"[185] Ultimately, in *AT&T*, Thomas was willing to read "any" in the statute contextually to permit only defenses to *formation* of contracts, not generally applicable contract law that applies to block enforcement of any contract whatsoever. By limiting the effect of "any" and focusing on contextual clues elsewhere in the text of § 2 and § 4, Thomas made a serious effort to vindicate a reading that would allow challenges to arbitration agreements arising from contract law defenses that apply only to block formation. Yet notwithstanding the textual and contextual arguments in his concurrence, he also provided the fifth vote for the majority's purposive analysis, albeit "reluctantly."[186]

The liberal dissenters doubled down on the textual "any" in a more literalist rendering of § 2.[187] But they did not ignore the centrality of purpose and the primary objective of the FAA in evaluating a statutory use of "any." Rather, citing relevant House and Senate Reports associated with the FAA, the dissent identified a narrower purpose to put arbitration agreements and any other agreements "upon the same footing"—reinforcing its reading of "any." Because the Senate Report specifically pointed to § 2 to express statutory purpose, the dissent thought the focus on streamlined procedures in the majority opinion was inappropriate as a purposive reading.[188] Thus, in the final analysis, the debate about "any" really turns on legislative purpose. Judges with different methodological commitments will surely use different kinds of evidence to divine purpose, but both the majority and the dissent should have been more forthright that the scope of "any" in the statute could only come from some argument about statutory purpose, however ascertained.

185. *AT&T*, 563 U.S. at 352 (Thomas, J., concurring).

186. *Id.* at 353.

187. *Id.* at 359 (Breyer, J., dissenting).

188. *Id.* at 359–62 (citing H.R. REP. NO. 96-68, at 1 (1924); S. REP. NO. 536-68, at 2 (1924)).

2. Reves v. Ernst & Young[189]

In this case, about the reach of the statutory language "any note" in the Securities Exchange Act of 1934,[190] the Court likewise interrogated congressional purpose to conclude that promissory notes payable on demand by holders did fall within the "any note" definitional language of "security" under the Act. Although there was a textualist partial dissent from a portion of Justice Marshall's otherwise unanimous opinion, the entire Court signed onto a purposive evaluation of Congress's intent in regulating securities broadly.

At issue for the Court was whether an auditor could be sued under the antifraud provisions of the Securities Exchange Act for inflating the assets and net worth of a farmers' cooperative that offered notes to members and nonmembers at variable interest rates in order to beat local financial institutions, and marketed them as an investment program backed by real assets held by the co-op and audited by the predecessor to Ernst & Young. Ernst & Young had successfully argued in the Eighth Circuit that the co-op notes could not be considered "any note" in the definition of "security" in the Act. But the Supreme Court disagreed.

Even before any textual analysis got off the ground, Justice Marshall invoked the "fundamental purpose undergirding the Securities Act[:] 'to eliminate serious abuses in a largely unregulated securities market.'"[191] He emphasized that "Congress painted with a broad brush[,] . . . recogniz[ing] the virtually limitless scope of human ingenuity, especially in the creation of 'countless and variable schemes devised by those who seek the use of the money of others on the promise of profits.'"[192] Although mostly citing the Court's own precedents to divine congressional purpose, Marshall also briefly quoted a House Report identifying that "the term 'security'" should be defined "in sufficiently broad and general terms."[193] He then openly acknowledged that in deciding which instruments the Court would deem "securities" under the Act, it was necessary not to employ "legal formalisms,

189. Reves v. Ernst & Young, 494 U.S. 56 (1990).

190. 15 U.S.C. § 78c(a)(10).

191. *Reves*, 494 U.S. at 60.

192. *Id.* at 60–61.

193. *Id.* at 61 (quoting H.R. REP. NO. 85-73, at 11 (1933)).

but instead [to] take account of the economics of the transaction," focusing on "Congress' purpose" in a pragmatic way to determine the ambit of the Act.[194]

All this led to the opinion being clear that "the phrase 'any note' should not be interpreted to mean literally 'any note,' but must be understood against the backdrop of what Congress was attempting to accomplish in enacting the Securities Acts."[195] Nevertheless, in light of that set of objectives the Court was willing to use the "any" language to establish a rebuttable "presumption that every note is a security."[196] From that presumption, the Court developed a test which drew upon lower court common law efforts to define which "notes" are "securities" under the Act and which are not—and the Court found no difficulty concluding that the co-op's notes were covered as securities, consistent with congressional purpose.[197]

Although the rest of the Court split in various ways about how to apply an exclusion at the end of the relevant statutory section about short-term maturity date notes—some offering more *stare decisis* and legislative history–focused arguments[198] and some offering more textual readings of the exception[199]—the whole Court (including partial dissenters Rehnquist, White, O'Connor and Scalia) joined in the purpose-driven construal of "any note" in the primary definition of the Act's coverage. Notwithstanding the partial dissenters' conventional use of an italicized "any" to argue for a broadened use of the short-term maturity exemption,[200] the separate opinion about the exclusion does nothing to undermine the appropriateness of drawing on congressional purpose to understand the primary definitional section of covered securities. Again, here, *Reves* supports the general use of congressional purposes, however divined, in arriving at contextually persuasive ways to read the word "any" in legislative work-product.

194. *Id.*
195. *Id.* at 63.
196. *Id.* at 65.
197. *Id.* at 67–70.
198. *Id.* at 73–76 (Stevens, J., concurring).
199. *Id.* at 76–82 (Rehnquist, J., concurring in part and dissenting in part).
200. *Id.* at 80.

IV. "ANY" AND CONTEMPORARY LEGISPRUDENCE

We don't mean for this inquiry into drafting and interpretive practices surrounding the word "any" to be just another screed against textualism—or another effort to reveal the limits of that approach for a Court dead-set on embracing it in one form or another. From our vantage point, it isn't simply that "ordinary meaning" textualists cannot agree on uses of "any" with the few tools they find relevant, defaulting too often to what they assert to be plain meanings. As we have shown above, these textualists have mostly relied on italicization, dictionaries, and some selectively invoked canons (linguistic, structural, and substantive) to bounce among purportedly clear meanings of "any" in any given statute. But we have also shown that they are capable, even within their textualisms, of reaching a little deeper for contextual cues to help inform their interpretations.

Although the Court's textualists haven't yet sought to use "ordinary meaning" textualism's newer gadgets—such as experimental jurisprudence[201] or corpus linguistics[202]—we doubt those technologies are going to solve the problem of reining in the discretion judges operate with when they decide how to read "any" in a particular statute. That is because construing "any" in a statutory setting requires developing a rich story about legislative context. Put most directly, "any" adjusts the meaning of *other* words; it has a *function* rather than a definition. Indeed, it can have multiple functions that can affect other words in a statute; and those functions or uses are rooted in the communicative circumstances of the statutory utterance. Even textualists who want to avoid intentionalist inquiries must recognize context to understand the role played by "any."

In life and law, universal quantifiers routinely are accompanied by a range of domain restrictions, furnished by background understandings that are appreciated by drafters and citizens alike. Yet rather than accepting that the Court should regularly look

201. *See* Tobia et al., *supra* note 16; Kevin Tobia, *Testing Ordinary Meaning*, 134 HARV. L. REV. 726 (2020).

202. *See generally* Thomas R. Lee & Stephen C. Mouritsen, *Judging Ordinary Meaning*, 127 YALE L.J. 788 (2018); Stephen C. Mouritsen, *Hard Cases and Hard Data: Assessing Corpus Linguistics as an Empirical Path to Plain Meaning,* 13 COLUM. SCI. & TECH. L. REV. 156 (2011); Lawrence M. Solan and Tammy Gales, *Corpus Linguistics as a Tool in Legal Interpretation*, 2017 BYU L. REV. 1311 (2018).

outside the dictionary and statutory structure, we too often get opinions in majority and dissent by textualists that argue about the scope of "any judgment regarding the granting of relief" (in a recent term, for example) without being willing to explore seriously congressional purpose and plan—context which could help resolve the dispute.[203]

We tried to show above that even committed textualists of various types have found ways to allow purposive readings of "any" to matter in some cases—and that seems a promising pathway to help the Court know when an expansive or limiting reading of "any" makes the most sense of the statutory scheme at issue. Perhaps some judges will read this Article and see the need for a new "any" canon that allows them to view almost all uses of "any" as requiring a fair look at a broad range of contextual resources. Although "ambiguity" is not really the right concept, because the literal meaning of "any" standing alone is rarely the best way to start the inquiry, more literalist judges will often need something akin to a threshold ambiguity determination to start looking broadly at context. Perhaps more judges will adopt Chief Justice Roberts' apparent willingness in statutory interpretation cases to permit a textualism inflected with careful attention to legislative planning and purpose.[204] Our larger point, however, applies outside of the textualism of the moment, for it isn't just conservative judges playing fast and loose with "any." Liberal judges too could be doing better in their "any" jurisprudence, as they often are content to play by seemingly textualist rules without mounting fuller contextual investigations.

In this regard, we share certain concerns expressed by other scholars about the hegemony of textualism. One is a doubt that the Court's persistent hyper-focus on statutory text makes these laws more understandable to ordinary citizens. As William Eskridge and Victoria Nourse put it, the growing inaccessibility of statutory meaning is due to a methodology that "combine[s] elasticity and opportunities for source-shopping with normative vacuity."[205]

---

203. Patel v. Garland, 142 S. Ct. 1614 (2022) (interpreting and quoting from 8 U.S.C § 1252(a)(2)(B)(i)).

204. *See* Bond v. United States, 572 U.S. 844 (2014); King v. Burwell, 576 U.S. 473 (2015). *See generally* Richard Re, *The New Holy Trinity,* 18 GREEN BAG (2D) 407 (2015); Abbe R. Gluck, *Imperfect Statutes, Imperfect Courts: Understanding Congress's Plan in the Era of Unorthodox Lawmaking*, 129 HARV. L. REV. 62, 64, 66 (2015).

205. *See* Eskridge & Nourse, *supra* note 15, at 1737.

The prototypical "average citizen" seeking to understand the domain of "any" in federal statutes will not readily comprehend why parallel sections of the FTCA, using "any" in closely similar settings, yield totally opposite resolutions based on divergent textualist approaches.[206] Or why the same language canon in one setting is deemed unnecessary to establish the textual clarity of "any," while in another case it establishes the scope of "any" as so textually clear that there is no need to consider legislative background.[207]

A related but distinct concern involves the declared objectivity of textualist sources, notably dictionaries and content-neutral language canons. We and others have suggested that this presumptively neutral or objective character is instrumentally valuable to those (including the justices) who seek to rebut charges of a politicized Court.[208] But the sheer number of dictionary definitions and the lack of any canonical hierarchy undermine notions that judges wielding these resources are sensibly constrained or are ceding policy choices to legislators. Our review of "any" decisions illustrates that this capacious statutory universe reflects something beyond definitions of words and interactive phrases. One important lesson is that when it comes to "any" in a statutory scheme, context is more than an external constraint on the word's scope; it is essential to determining that word's use or function in the scheme.[209]

More generally, statutes consist of substantive rules and standards established and negotiated by legislators, who draft and approve those rules and standards in order to accomplish certain objectives or purposes. When construing "any"—whether "labor or service of any kind," "any tangible object," "any other law enforcement officer," or "convicted in any court," to use examples we have discussed in this Article—a respect and appreciation for those purposes would seem essential if we are to be governed by

206. *See supra* Part III.A (comparing divergent readings of "any" in the FTCA in *Dolan* and *Ali*).

207. *Compare supra* Part III.B.1 (discussing the rejection of *ejusdem generis* in *Harrison*) *with* Part III.B.2 (discussing the preclusive impact of *ejusdem generis* in *Circuit City*).

208. *See* James J. Brudney & Corey Ditslear, *Canons of Construction and the Elusive Quest for Neutral Reasoning*, 58 VAND. L. REV. 1, 103–05 (2005) (re: canons); Brudney & Baum, *supra* note 40, at 499–500 (re: dictionary reliance); Margret H. Lemos, *The Politics of Statutory Interpretation*, 89 NOTRE DAME L. REV. 849, 885–86 (2014) (book review) (discussing the apparent "rulinesss" of textualism).

209. *Cf.* SLOCUM, *supra* note 17, at 167.

legislative rather than judicial policy determinations. Even in the cases in which substantive canons inform the domain of "any," at least some of those canons emerge over time as a background against which many legislators govern.[210]

As we have shown in Part II, isolating bits of text such as "any" and tearing "any" away from a richer contextual analysis of a statutory scheme is not just a textualist misstep by conservative judges. Rather, the dynamic of "textual isolationism" that Nourse has identified in many Court decisions[211] is—while not an equal opportunity offense among liberal judges—something justices on the left similarly find themselves doing too when it comes to "any." It just won't do for judges to isolate "any" from a narrative of what the statute is trying to accomplish in its coverage or exceptions; textualists and purposivists alike need to remember as much.

If this Article accomplishes nothing else, we hope it invites all judges to embrace greater humility when adjudicating cases that turn on "any." We know textualists won't suddenly support a new "information economy," as Nourse puts it,[212] to increase their likelihood of getting "any" right with legislative history. But we also know all judges—right-wing and left-wing, textualist and purposivist—can do better by admitting reflections about a statute's plan or scheme when deciding what "any" is doing in context and thinking about how "any" interfaces with other normative commitments evident in a set of well-entrenched linguistic and substantive canons. Although Nourse is modestly more optimistic that canons like *ejusdem generis* can help as a small corrective for isolationism,[213] our study of "any" reveals that whether it is appropriate to apply *ejusdem generis*—or for that matter policy presumptions against preemption or extraterritorial application—probably requires a threshold investigation into the legislative plan or scheme. That is part of another takeaway from our inquiry into

210. There is mixed evidence about the ways the federalism and lenity canons may or may not make their way into drafting practices. *See* Gluck & Bressman, *supra* note 35, at 941–47 (exploring congressional staffers' knowledge of some substantive canons).

211. *See* Nourse, *supra* note 14. Nourse accuses purposivists of this practice too, to be clear, and our survey of judges in their "any" decisions confirms the point. She also analyzes *Yates* in her paper—as we have above at various junctures in the Article—but she only highlights the complexity associated with "any" in that case in a footnote. *See id.* at 1422 n.72.

212. *Id.* at 1413.

213. *Id.* at 1433 ("In this sense, [*ejusdem generis*] should be seen as a defensive, rather than offensive tool, resisting isolationism . . . .").

"any": some words especially trigger the need for illuminating extrinsic context, and using only intrinsic sources of meaning limits judicial wisdom while increasing the risks of unbridled judicial discretion.

CONCLUSION

Our effort to expose variable drafting and interpretive approaches to the ubiquitous word "any"—dispersed throughout our statute books and federal reporters—is, in part, a case study about the way a word which has multiple uses needs a lot of context for sensitive interpretation appropriate to the relevant statutory setting. Literalism just won't do—as philosophers and linguists recognize regarding universal quantifiers[214] and as citizens know through their common sense. This Article—which examined legislative drafting manuals, surveyed centuries of Court decisions, and conducted in-depth pairwise comparisons of "any" cases—reveals that struggles in construing "any" come with the territory of dealing with a legislatively valued word that can do so many different things in statutory provisions. Our recommendation for the Court is that its "any" jurisprudence would be improved by conceding that "any" is facially inconclusive most of the time—and then permitting itself the widest range of contextual sources to decide its scope and function: call this a new "any" canon. We aren't expecting the modal textualist to admit legislative history but we have shown that, notwithstanding frequent pronouncements about the "plain meaning" of the word, textualists and non-textualists alike have been willing to interrogate the legislative record and policy presumptions in service of purposive readings. Integrity requires judges to acknowledge that interpreting "any" will require a panoply of sources to inform its use in a given statutory context.

214. *See generally* SLOCUM, *supra* note 17; Stanley & Szabó, *supra* note 18.

Bill of Rights Nondelegation

Eli Nachmany[*]

Speculation about the "revival" of the nondelegation doctrine has reached a fever pitch. Although the Supreme Court apparently has not applied the nondelegation doctrine to declare a federal statute unconstitutional since 1935, the doctrine may be making a comeback. The common understanding is that the nondelegation doctrine prohibits Congress from "delegating" legislative power to the executive branch. While the nondelegation doctrine may appear to be about limiting Congress, its ultimate target is delegation. But if the nondelegation doctrine is about policing delegation, then the Court has been regularly—and rigorously—applying the doctrine in a different context: In litigation concerning various provisions of the Bill of Rights, the Court has enforced a nondelegation principle to constrain the delegation of unfettered discretion to the executive.

The uncovering of a Bill of Rights nondelegation doctrine reveals that, contrary to popular belief, the Court has been actively applying some form of nondelegation for many years. Recognizing a Bill of Rights nondelegation doctrine could have important implications for Bill of Rights jurisprudence writ large. Further, understanding the "Bill of Rights nondelegation doctrine" as a coherent line of cases separate from what this Article calls the "Article I nondelegation doctrine" helps to clarify the connection that some have pointed out between the nondelegation principle and certain parts of the Bill of Rights. From the First and Second

* Law clerk to Judge Steven J. Menashi, U.S. Court of Appeals for the Second Circuit, 2022–2023. Harvard Law School, J.D. 2022; New York University, B.S. 2017. I am grateful to Erin Blondel, Sam Bray, Louis Capozzi, John Duffy, Richard Fallon, Noah Feldman, Charles Fried, Kristin Hickman, Orin Kerr, Christa Laser, Gary Lawson, Judge Menashi, Alexandra Natapoff, Aaron Nielson, Mark Shawhan, Ben Silver, Lawrence Solum, Chad Squitieri, Adrian Vermeule, Adam White, and Ilan Wurman for helpful comments on earlier drafts of this Article. This Article also benefitted from feedback received during presentations at the C. Boyden Gray Center for the Study of the Administrative State at the George Mason University Antonin Scalia Law School and the Federalist Society's Junior Scholars Colloquium. All errors are my own.

Amendments to the Fourth and Fifth Amendments, the Bill of Rights nondelegation doctrine prevents the delegation of unfettered discretion when enumerated rights are at stake.

CONTENTS

INTRODUCTION

The familiar conception of the nondelegation doctrine is something like the following: Congress cannot delegate its legislative power to the executive branch.[1] In practice, if Congress passes a statute that would permit the executive to exercise legislative power in carrying out the law, a court applying the nondelegation doctrine would declare the offending statutory provision unconstitutional.[2] Scholars and jurists disagree on whether the nondelegation doctrine is consistent with the original meaning of the Constitution,[3] serves constitutional values,[4] or even exists.[5] The debates rage on. But against the backdrop of serious scholarly and jurisprudential inquiry into the doctrine's propriety, all would likely agree that the Supreme Court has signaled its openness in recent years to "reviving" the nondelegation doctrine.[6]

1. *See* Thomas W. Merrill, *Rethinking Article I, Section 1: From Nondelegation to Exclusive Delegation*, 104 COLUM. L. REV. 2097, 2098–99 (2004). To be sure, the nondelegation doctrine is not only about Congress and the Executive. The nondelegation doctrine also prohibits Congress from delegating legislative power to the judicial branch. *See* Wayman v. Southard, 23 U.S. (10 Wheat.) 1 (1825); *see also* Aaron Nielson, Erie *as Nondelegation*, 72 OHIO ST. L.J. 239, 266 (2011) (noting the focus in the scholarly literature on delegations to the Executive as opposed to delegations to courts).

2. *See, e.g.*, A.L.A. Schechter Poultry Corp. v. United States, 295 U.S. 495 (1935) (declaring section 3 of the National Industrial Recovery Act of 1933 unconstitutional on nondelegation grounds).

3. *Compare* Julian Davis Mortenson & Nicholas Bagley, *Delegation at the Founding*, 121 COLUM. L. REV. 277 (2021), *with* Ilan Wurman, *Nondelegation at the Founding*, 130 YALE L.J. 1490 (2021). *See also* Jennifer Mascott, *Early Customs Laws and Delegation*, 87 GEO. WASH. L. REV. 1388 (2019).

4. *Compare* Gundy v. United States, 139 S. Ct. 2116, 2130 (2019), *with* David Schoenbrod, *Consent of the Governed: A Constitutional Norm That the Court Should Substantially Enforce*, 43 HARV. J.L. & PUB. POL'Y 213 (2020), *and Gundy*, 139 S. Ct. at 2131 (Gorsuch, J., dissenting).

5. *Compare* Keith E. Whittington & Jason Iuliano, *The Myth of the Nondelegation Doctrine*, 165 U. PA. L. REV. 379 (2017), *and* Eric A. Posner & Adrian Vermeule, *Interring the Nondelegation Doctrine*, 69 U. CHI. L. REV. 1721 (2002), *with* Douglas H. Ginsburg, *Reviving the Nondelegation Principle in the US Constitution*, *in* THE ADMINISTRATIVE STATE BEFORE THE SUPREME COURT: PERSPECTIVES ON THE NONDELEGATION DOCTRINE 20 (Peter J. Wallison & John Yoo eds., 2022) [hereinafter PERSPECTIVES ON NONDELEGATION], *and* Kristin E. Hickman, *Nondelegation as Constitutional Symbolism*, 89 GEO. WASH. L. REV. 1079 (2021), *and* Aaron Gordon, *Nondelegation*, 12 N.Y.U. J.L. & LIBERTY 718 (2019), *and* Larry Alexander & Saikrishna Prakash, *Reports of the Nondelegation Doctrine's Death Are Greatly Exaggerated*, 70 U. CHI. L. REV. 1297 (2003).

6. In *Gundy v. United States*, Justice Gorsuch dissented from the Court's decision to uphold a sex offender registration statute that had been challenged on nondelegation grounds. The statute—known by the acronym SORNA (Sex Offender Registration and

"Reviving" is in scare quotes for a reason. The belief that application of the nondelegation doctrine constitutes a "revival" proceeds from the view that the Court has not applied it in nearly a century.[7] The traditional understanding of the nondelegation doctrine is that it flows from the Vesting Clause of Article I of the Constitution,[8] preventing Congress from delegating a vested legislative power. To that end, this Article will refer to this version of the nondelegation doctrine as the "Article I nondelegation doctrine." And the traditional view is correct—the Court has not applied *this version* of the nondelegation doctrine to hold a statute unconstitutional since 1935. That year, the Court declared the National Industrial Recovery Act of 1933—a centerpiece of President Franklin Delano Roosevelt's New Deal agenda—an unconstitutional delegation of legislative power.[9] In 2000, Cass Sunstein wrote that "the conventional doctrine has had one good year, and 211 bad ones (and counting)."[10] Over two decades later, one might say that the nondelegation doctrine still has not had another "good year."

Yet the Court's Bill of Rights jurisprudence indicates that the nondelegation doctrine is alive and well. In fact, the Court has regularly—and rigorously—applied a form of the nondelegation

Notification Act)—granted authority to the Attorney General of the United States to "'specify the applicability' of SORNA's registration requirements" for "individuals convicted of a sex offense before SORNA's enactment." *Gundy*, 139 S. Ct. at 2122 (majority opinion). Justice Gorsuch articulated a strong conception of the nondelegation doctrine, and Chief Justice Roberts and Justice Thomas joined Justice Gorsuch's dissenting opinion. *See id.* at 2131, 2133–37 (Gorsuch, J., dissenting). Justice Alito authored a concurring opinion that signaled possible support for the doctrine while articulating that he felt its application would be inappropriate in the instant case "because a majority [was] not willing to do that[.]" *See id.* at 2130–31 (Alito, J., concurring in the judgment). Justice Kavanaugh joined the Court after *Gundy* was argued, and in a statement respecting the denial of certiorari in a case similar to *Gundy*, he expressed measured support for the nondelegation doctrine as well. *See* Paul v. United States, 140 S. Ct. 342, 342 (2019) (Kavanaugh, J., statement respecting the denial of certiorari) ("Justice Gorsuch's thoughtful *Gundy* opinion raised important points that may warrant further consideration in future cases.").

7. For an example of the view that application of the nondelegation doctrine would constitute a revival, see, for example, Mike Rappaport, *Reviving the Nondelegation Doctrine*, L. & LIBERTY (Mar. 8, 2018), https://lawliberty.org/reviving-the-nondelegation-doctrine.

8. *See* Cary Coglianese, *Six Degrees of Delegation*, REGUL. REV. (Dec. 23, 2019), https://www.theregreview.org/2019/12/23/coglianese-six-degrees-delegation; *see also* U.S. Const. art. I, § 1 ("All legislative Powers herein granted shall be vested in a Congress of the United States, which shall consist of a Senate and House of Representatives.").

9. *See* A.L.A. Schechter Poultry Corp. v. United States, 295 U.S. 495 (1935).

10. Cass R. Sunstein, *Nondelegation Canons*, 67 U. CHI. L. REV. 315, 322 (2000).

doctrine to enforce various provisions of the Bill of Rights. This Article reveals the existence of a Bill of Rights nondelegation doctrine. In Bill of Rights litigation, the Court has developed something like a nondelegation doctrine to evaluate whether certain infringements on individual liberty are permissible.[11] Here, the Court disfavors the delegation of discretion.

At bottom, the nondelegation doctrine is about policing *delegation*. The prohibition on Congress delegating legislative power to the executive is, in fact, a prohibition on Congress delegating to the executive the ability to exercise a kind of discretionary power pursuant to the executive's own will. Importantly, the Article I nondelegation doctrine is not a *substantive* limit on Congress's power; it is a *procedural* one. Suppose that Congress passes a law that significantly constrains individual liberty, but it sets forth detailed instructions for how the executive is to carry out the law. One might challenge the law on the grounds that Congress has exceeded its own legislative power,[12] but a nondelegation objection would fail. So long as Congress has made the relevant policy choices, it has not *delegated* legislative power.

The Bill of Rights nondelegation doctrine cuts a similar profile. The Supreme Court's Bill of Rights jurisprudence has recognized procedural limits on delegation with respect to several amendments. The Court sees the Constitution as allocating power between branches of government, and when certain of the Bill of Rights amendments are at issue, discretion may not be delegated in a way that upsets that allocation. In First Amendment cases, for example, the Court has long taken the view that discretionary permitting regimes for speech are themselves censorious and thus unconstitutional.[13] The Court feels similarly about the Second

11. This Article takes no position on whether the Bill of Rights nondelegation doctrine is consistent with the original meaning of the Constitution.

12. For example, one might challenge an economic regulatory statute on the ground that it exceeds Congress's power "[t]o regulate Commerce . . . among the several States"—one of the substantive legislative powers that the Constitution vests in Congress. U.S. CONST. art. I, § 8, cl. 3.

13. City of Lakewood v. Plain Dealer Publ'g Co., 486 U.S. 750, 757 (1988) ("[A] licensing statute placing unbridled discretion in the hands of a government official or agency constitutes a prior restraint and may result in censorship." (first citing Shuttlesworth v. Birmingham, 394 U.S. 147, 151 (1969); then citing Cox v. Louisiana, 379 U.S. 536 (1965); then citing Staub v. City of Baxley, 355 U.S. 313, 321–22 (1958); then citing Kunz v. New York, 340 U.S. 290, 294 (1951); then citing Niemotko v. Maryland, 340 U.S. 268 (1951); and then citing Saia v. New York, 334 U.S. 558 (1948))).

Amendment, as its recent decision in *New York State Rifle & Pistol Association v. Bruen*[14]—and Justice Kavanaugh's concurrence in the majority's opinion[15]—demonstrates.[16] But the Bill of Rights nondelegation doctrine is not only about permitting regimes for speech and guns. Fourth Amendment cases concerning the right's particularity requirement have keyed in on the delegation from courts to the executive that occurs when judges approve warrants that permit police to exercise too much discretion.[17] Moreover, the Fifth Amendment's (and Fourteenth Amendment's) void-for-vagueness doctrine—as the Supreme Court has applied it—has prohibited the legislature from delegating penal lawmaking discretion to the executive.[18]

This Article rethinks the conventional wisdom on nondelegation. It reveals the existence of a nondelegation doctrine in a line of cases wholly separate from the Article I doctrine that dominates much of the scholarly literature on nondelegation. The connection between the nondelegation doctrine and discretion is well-known.[19] And some have even drawn the parallel between the void-for-vagueness doctrine and the nondelegation doctrine.[20] But this Article is the first piece of scholarship explicitly tying together—and shining a light on—a coherent nondelegation doctrine for the Bill of Rights.

Some caveats are necessary. Most importantly, this Article does not take the position that the Bill of Rights nondelegation doctrine and the Article I nondelegation doctrine are perfectly analogous.

14. N.Y. State Rifle & Pistol Ass'n v. Bruen, 142 S. Ct. 2111 (2022).

15. *Id.* at 2161 (Kavanaugh, J., concurring) (describing New York's "unusual discretionary licensing regime[]" for granting concealed-carry permits as unconstitutional).

16. *Id.* (describing the "open-ended discretion" that New York licensing officials enjoyed); *see also id.* at 2123 (majority opinion) (contrasting New York's regime with those of states in which "authorities *must* issue concealed-carry licenses whenever applicants satisfy certain threshold requirements" (emphasis added)).

17. *See, e.g.*, United States v. Leon, 468 U.S. 897, 923 (1984).

18. *See* Johnson v. United States, 576 U.S. 591, 595 (2015); *see also* Sessions v. Dimaya, 138 S. Ct. 1204 (2018) (applying the void-for-vagueness doctrine outside of the criminal context to civil deportation).

19. *See, e.g.*, Aditya Bamzai, *Delegation and Interpretive Discretion:* Gundy, Kisor, *and the Formation and Future of Administrative Law*, 133 HARV. L. REV. 164 (2019); Gary Lawson, *Discretion as Delegation: The "Proper" Understanding of the Nondelegation Doctrine*, 73 GEO. WASH. L. REV. 235 (2005); *see also* Ilan Wurman, *As-Applied Nondelegation*, 96 TEX. L. REV. 975, 980 (2018) (describing the "'intelligible principle' test" as "entirely a question of discretion").

20. *See, e.g.*, Todd Gaziano & Ethan Blevins, *The Nondelegation Test Hiding in Plain Sight: The Void-for-Vagueness Standard Gets the Job Done*, *in* PERSPECTIVES ON NONDELEGATION, *supra* note 5, at 45; Arjun Ogale, Note, *Vagueness and Nondelegation*, 108 VA. L. REV. 783 (2022).

Moreover, the Bill of Rights nondelegation doctrine does not have anything to say about the executive's inherent prosecutorial discretion, which does not come to the executive by way of delegation. Furthermore, the Bill of Rights nondelegation doctrine is not the only way that the Court enforces the various protections of the Bill of Rights; it is merely one of a few tools in the Court's rights-protective toolbox. Nor does the "Bill of Rights nondelegation doctrine" necessarily apply to each provision of the Bill of Rights. That said, the Bill of Rights nondelegation cases offer important insights about how the Court enforces the guarantees of the Bill of Rights and applies the nondelegation principle.

This Article proceeds in four parts. Part I provides background on the Article I nondelegation doctrine's history. Part II uncovers a separate Bill of Rights nondelegation doctrine, tying together the relevant First, Second, Fourth, and Fifth Amendment cases into a coherent framework. Part III considers some potential applications of the Bill of Rights nondelegation doctrine. Part IV then distinguishes the Bill of Rights nondelegation doctrine from the Article I nondelegation doctrine, while articulating other important caveats to the parallels drawn in this Article. Part IV also locates the Bill of Rights nondelegation doctrine within the broader framework of the protections of the Bill of Rights.

Certainly, the Article I nondelegation doctrine has only had "one good year" since 1935. But the Bill of Rights nondelegation doctrine has had plenty of good years. As this Article will reveal, the doctrine is firmly ensconced in the Court's jurisprudence.

I. THE ARTICLE I NONDELEGATION DOCTRINE

To understand the Bill of Rights nondelegation doctrine, one must understand how it is distinct from a wholly separate doctrine of nondelegation: the Article I nondelegation doctrine. This Part surveys the development of this doctrine in American constitutional law. When scholars and jurists speak of the "nondelegation doctrine," often what they are talking about is the Article I nondelegation doctrine.[21] To demonstrate the point, this Part locates the textual and structural sources of the Article I

21. *But see* Benjamin Silver, *Nondelegation in the States*, 75 VAND. L. REV. 1211, 1216 (2022) ("State nondelegation law arises in a wide variety of contexts, not simply the legislature-to-agency contexts with which scholars and federal courts are most familiar.").

nondelegation doctrine. From there, this Part surveys the doctrine's development in the Supreme Court.

A. Locating the Constitutional Source of the Article I Nondelegation Doctrine

Various scholars and jurists take the position that Congress may not delegate any of its vested legislative powers—in whole or in part—to another branch of the federal government. Often, the purported textual sources of this principle of nondelegation are the Vesting Clause of Article I and the structure of the Constitution (thus, this Article refers to this version of the doctrine as the "Article I nondelegation doctrine"). In the interest of supporting the claim that the Article I nondelegation doctrine exists, nondelegation proponents must argue that the Constitution—properly understood—contains this principle. This section explores the argument that the Constitution contains a nondelegation doctrine in Article I.

1. Article I Nondelegation: An Argument from Text and Structure

The Constitution separates power. While the dividing lines of this separation are perhaps not entirely clean,[22] the fact remains that the separation of powers is one of the core organizing principles—if not *the* central principle—of the Constitution's framework.[23] To that end, Articles I, II, and III of the Constitution lay out a structure of government in which (as a general matter) three different institutions respectively exercise three different sorts of power.

22. *See* THE FEDERALIST NO. 47 (James Madison); *see also, e.g.*, Andrew Kent & Julian Davis Mortenson, *Executive Power and National Security Power*, *in* THE CAMBRIDGE COMPANION TO THE UNITED STATES CONSTITUTION 261–91 (Karen Orren & John W. Compton eds., 2018) (demonstrating tension between the defining role of the executive as law-executor and presidential power in the national security realm); Frederick Green, *Separation of Governmental Powers*, 29 YALE L.J. 369, 384 n.60 (1920) ("The pardoning power, like the veto, is a legislative power of negative nature, vested by the constitution in the chief executive . . . ."). One might also say that the Senate's power to withhold consent to (and thereby block) certain of the president's nominees for positions in the executive branch—a power housed in Article II of the U.S. Constitution, *see* U.S. CONST. art. II, § 2, cl. 2—is a sort of executive power reposed in (part of) the legislative branch.

23. *See* THE FEDERALIST NO. 51, at 75 (James Madison) (Benediction Classics 2016); *see also Considering the Role of Judges Under the Constitution of the United States: Hearing Before the S. Comm. on the Judiciary*, 112th Cong. 6–8 (2011) (statement of Justice Antonin Scalia) (venerating the separation of powers as safeguarding individual liberty and the protections of the Bill of Rights).

In particular, Article I sets up the legislative branch, Article II sets up the executive branch, and Article III sets up the judicial branch.[24]

Each Article begins with what is called a "Vesting Clause," vesting power in the branch of government that the Article establishes. The three Articles begin similarly, but the first clause of Article I contains an important linguistic difference from the first clauses of Articles II and III. Compare the following three opening clauses:

- Article I: "All legislative powers herein granted shall be vested in a Congress of the United States, which shall consist of a Senate and House of Representatives."[25]
- Article II: "The executive power shall be vested in a President of the United States of America."[26]
- Article III: "The judicial power of the United States, shall be vested in one supreme Court, and in such inferior Courts as the Congress may from time to time ordain and establish."[27]

Unlike the Article II and III Vesting Clauses, Article I's Vesting Clause begins with the word "All." In this way, the Article I Vesting Clause does not vest in Congress a free-floating federal legislative power. Rather, it vests in Congress no more (and no less) than the legislative powers "herein granted."

To some extent, the Article I Vesting Clause is a linguistic minefield of interpretation. Uncertainty about the meaning of various terms in the Article I Vesting Clause—including "legislative" powers,[28] "herein granted,"[29] and "vested"[30]—has provided significant fodder for scholarly inquiry in recent years. This scholarship has introduced nuance into the task of interpreting the Article I Vesting Clause. Nevertheless, the classic understanding

24. *See* U.S. CONST. art. I, § 1; *id.* art. II, § 1, cl. 1; *id.* art. III, § 1.

25. U.S. CONST. art. I, § 1.

26. U.S. CONST. art. II, § 1, cl. 1.

27. U.S. CONST. art. III, § 1.

28. *See, e.g.*, Mortenson & Bagley, *supra* note 3, at 294–95; Neomi Rao, *Why Congress Matters: The Collective Congress in the Structural Constitution*, 70 FLA. L. REV. 1, 9–21 (2018).

29. *See, e.g.*, Richard Primus, *Herein of "Herein Granted": Why Article I's Vesting Clause Does Not Support the Doctrine of Enumerated Powers*, 35 CONST. COMMENT. 301, 302–03, 302 n.6 and accompanying text (2020); Coglianese, *supra* note 8.

30. *See, e.g.*, Jed Shugerman, *Vesting*, 74 STAN. L. REV. 1479 (2022); Mortenson & Bagley, *supra* note 3, at 309–10; Gary Lawson, *Delegation and Original Meaning*, 88 VA. L. REV. 327, 337 (2002) [hereinafter Lawson, *Delegation and Original Meaning*].

of this constitutional provision begins from the following premise: Article I vests in Congress all of the legislative powers set forth in Article I of the Constitution. From there, proponents of the nondelegation doctrine argue that because *the people* have vested these legislative powers in Congress, the Constitution forbids Congress from *delegating* any of these powers to another branch. In other words, if the federal government is going to legislate pursuant to one of the legislative powers in Article I, Congress must be the one to do it. Congress cannot authorize another branch—and in practice, this often means the executive branch[31]—to do so.

Earlier conceptions of the Article I nondelegation doctrine located the doctrine's constitutional source in Article I's Vesting Clause.[32] But recent scholarship suggests that the source of the doctrine may be a combination of both the Vesting Clause *and* the general structure of the Constitution (or even just the latter).[33] The two are related. The Constitution sets up a structure of government that vests different sorts of power in different branches—the executive power in the executive branch, the judicial power in the judicial branch, and a certain reservoir of legislative powers in the legislative branch. One might even say that "[i]f Congress could pass off its legislative power to the executive branch, the '[v]esting [c]lauses, and indeed the entire structure of the Constitution,' would 'make no sense.'"[34]

The purpose of this Article is not to enter the Article I nondelegation debates; it takes no position on whether nondelegation proponents are correct that the Article I nondelegation doctrine comports with the original meaning of the Constitution. Scholars have spilled a significant amount of ink on the question,[35] and this Article's central insight does not rise or fall on the debate's resolution.

31. *But cf.* Mistretta v. United States, 488 U.S. 361 (1989) (reviewing a delegation to the judicial branch).

32. *See, e.g.*, A.L.A. Schechter Poultry Corp. v. United States, 295 U.S. 495, 529–30 (1935).

33. *See* Mascott, *supra* note 3, at 1395; *see also* Gundy v. United States, 139 S. Ct. 2116, 2134–35 (2019) (Gorsuch, J., dissenting) (quoting Lawson, *Delegation and Original Meaning*, *supra* note 30, at 340).

34. *Gundy*, 139 S. Ct. at 2134–35 (second and third alternation in original) (quoting Lawson, *Delegation and Original Meaning*, *supra* note 30, at 340); *cf.* City of St. Louis v. Dep't of Transp., 936 F.2d 1528, 1534 (8th Cir. 1991) ("No one claims, incidentally, that the delegation here was so broad as to violate Article I.") (explicitly grounding the nondelegation doctrine in Article I).

35. *See* Mascott, *supra* note 3.

Rather, the point of this subpart is that one can conceivably recognize a version of the nondelegation doctrine in Article I's Vesting Clause and the Constitution's general structure. The traditional understanding of that doctrine, discussed further in section I.B, sets the stage for the uncovering of a Bill of Rights nondelegation doctrine.

B. "One Good Year": The Article I Nondelegation Doctrine in Practice

The Article I nondelegation doctrine is a doctrine of judicial review. It is also a doctrine that the Supreme Court has rarely used. Indeed, the Court has only declared a federal statute unconstitutional on Article I nondelegation grounds twice (both times in 1935). Granted, the Court recognized a nondelegation doctrine in Article I as early as 1825. But employment of the Article I nondelegation principle in the 1930s was one of the main catalysts for a dark period in the Court's history—President Franklin Delano Roosevelt's proposal to "pack" the Court with Justices more sympathetic to his New Deal economic program. The history of President Roosevelt's court-packing proposal potentially provides a concrete explanation for why the Court backed away from using the Article I nondelegation doctrine in the exercise of judicial review. In the wake of the court-packing proposal, the Court has continued to acknowledge the existence of the Article I nondelegation doctrine, but it has not since applied the doctrine to declare a federal statute unconstitutional. That said, recent developments suggest that the Article I nondelegation doctrine may be making a return.

1. Article I Nondelegation Before the New Deal

In the words of Keith Whittington and Jason Iuliano, an examination of "the pre-New Deal tradition of [Article I] nondelegation jurisprudence . . . reveals that the constitutional limitation on the delegation of legislative power was frequently observed in theory but rarely enforced in practice."[36] This tradition began in the early 1800s, when the Marshall Court "heard the earliest cases challenging the unconstitutional delegation of

36. Whittington & Iuliano, *supra* note 5, at 383.

legislative power."[37] For the first quarter of the nineteenth century, the Court did not apply a particular rule of nondelegation.[38]

But in an 1825 case—*Wayman v. Southard*[39]—Chief Justice Marshall kicked off the Article I nondelegation doctrine's development at the Court. In *Wayman*, the Court stated an important principle of law: "It will not be contended that Congress can delegate to the Courts, or to any other tribunals, powers which are strictly and exclusively legislative. But Congress may certainly delegate to others, powers which the legislature may rightfully exercise itself."[40] The next paragraph began: "The line has not been exactly drawn which separates those important subjects, which must be entirely regulated by the legislature itself, from those of less interest, in which a general provision may be made, and power given to those who are to act under such general provisions to fill up the details."[41] The underlying delegation of interest to the Court in *Wayman* was a statutory provision in the Judiciary Act of 1789 that gave federal courts the power to regulate their own civil procedure.[42] But the idea that certain delegations are okay and others are not okay was a constitutionally significant proposition—necessarily, its corollary was that courts had the power to declare certain laws unconstitutional for effectuating an impermissible delegation of "powers which are strictly and exclusively legislative."[43]

Chief Justice Marshall's dicta in *Wayman* seemed to establish the ground rules for the Article I nondelegation doctrine. Congress could, in some instances, delegate power to a coordinate branch of the federal government. The big question seemed to be where one drew the line between "those important subjects" and "those of less interest."[44] Gary Lawson has suggested that the test for delegations might simply be that "Congress must make whatever decisions are

37. Keith E. Whittington, *Judicial Review of Congress Before the Civil War*, 97 GEO. L.J. 1257, 1291 (2009).

38. *See* Whittington & Iuliano, *supra* note 5, at 392–94; Andrew J. Ziaja, *Hot Oil and Hot Air: The Development of the Nondelegation Doctrine Through the New Deal, a History, 1813–1944*, 35 HASTINGS CONST. L.Q. 921, 925–28 (2008).

39. Wayman v. Southard, 23 U.S. (10 Wheat.) 1 (1825).

40. *Id.* at 42–43.

41. *Id.* at 43.

42. *See id.* at 43 (citing sections 7 and 17 of the Judiciary Act of 1789). We know these rules today as the Federal Rules of Civil Procedure.

43. *Id.* at 42.

44. *Id.* at 43.

important enough to the statutory scheme in question so that Congress must make them."[45] But perhaps this formulation is not exactly what Chief Justice Marshall was getting at. Rather, *Wayman*'s dicta may stand for the notion that certain congressional *powers* (like the power to regulate interstate commerce) are important—and nondelegable—while other *powers* (like the power to establish post offices) are "of less interest" and thus susceptible of delegation.[46]

A little over a century later, the Court articulated an authoritative test for the Article I nondelegation doctrine. In *J.W. Hampton, Jr. & Co. v. United States*,[47] the Court reviewed the constitutionality of a congressional grant of tariff-adjusting power to the President.[48] The Court upheld the statute against a nondelegation challenge, citing the fact that Congress had established a clear "policy and plan" for how the President was to carry out the law.[49] Explaining its reasoning, the Court stated that "[i]f Congress shall lay down by legislative act an *intelligible principle* to which the person or body authorized to fix such rates is directed to conform, such legislative action is not a forbidden delegation of legislative power."[50] *J.W. Hampton*'s rule represented a bit of a drift from the earlier, more muscular conceptions of the nondelegation doctrine.[51] The intelligible-principle test was forgiving; it permitted some delegation so long as Congress prescribed a standard by which a court could measure the executive's compliance with a given statute's command.

45. Lawson, *Delegation and Original Meaning*, *supra* note 30, at 361. Lawson's view of *Wayman* appears to have evolved in recent years. *See* Gary Lawson, *A Private-Law Framework for Subdelegation*, *in* PERSPECTIVES ON NONDELEGATION, *supra* note 5, at 123.

46. A.J. Kritikos, *Resuscitating the Non-Delegation Doctrine: A Compromise and an Experiment*, 82 MO. L. REV. 441, 446 (2017) (presenting this view). Another scholar has proposed "that the nondelegation doctrine be transformed into a series of nondelegation doctrines, each corresponding to one of Congress' distinct powers." Chad Squitieri, *Towards Nondelegation Doctrines*, 86 MO. L. REV. 1239, 1239 (2021).

47. J.W. Hampton, Jr., & Co. v. United States, 276 U.S. 394 (1928).

48. *See id.* at 404.

49. *See id.* at 405, 410–11.

50. *Id.* at 409 (emphasis added).

51. In an important nondelegation article, Ben Silver has shown that "the 'intelligible principle' test . . . was crafted under the influence of state nondelegation law." Silver, *supra* note 21, at 1260.

2. The Article I Nondelegation Doctrine Meets the New Deal and the Court-Packing Plan

If *J.W. Hampton* signaled a more permissive approach in the Court's nondelegation jurisprudence, that shift was not immediately felt. By the mid-1930s, America was in the throes of an economic depression.[52] In response, President Franklin Delano Roosevelt pushed Congress to enact his "New Deal" economic program.[53] "One of the cornerstones of the New Deal was the National Industrial Recovery Act of 1933 (NIRA)[,]" which gave the Roosevelt administration wide-ranging discretion to manage the economy as the United States grappled with financial calamity.[54] For example, the NIRA authorized the Roosevelt administration to enact "codes of fair competition[.]"[55] The Roosevelt administration proceeded to regulate with a heavy hand. In response, impacted businesses sought recourse in the federal courts, challenging the constitutionality of key aspects of the New Deal.[56] These challenges would lead the Court to apply the Article I nondelegation doctrine.

Perhaps the most famous nondelegation case from this era is *A.L.A. Schechter Poultry Corp. v. United States*.[57] The case stemmed from the indictment of Jewish poultry slaughterhouse operators in New York for violations of the "Live Poultry Code": a series of poultry regulations promulgated by the Roosevelt administration pursuant to the power conferred under the NIRA.[58] The NIRA provided that after "one or more trade or industrial associations or groups" submitted an application to the President, he could "approve a code or codes of fair competition for the trade or industry or subdivision thereof, represented by the applicant or applicants."[59] And on April 13, 1934, President Roosevelt approved

52. *See President Franklin Delano Roosevelt and the New Deal*, LIBR. OF CONG., https://www.loc.gov/classroom-materials/united-states-history-primary-source-timeline/great-depression-and-world-war-ii-1929-1945/franklin-delano-roosevelt-and-the-new-deal (last visited Oct. 29, 2023).

53. *See id.*

54. *See* Ziaja, *supra* note 38, at 942.

55. *Id.*

56. *See id.* at 943, 951 (describing businesses that brought suit against the federal government).

57. A.L.A. Schechter Poultry Corp. v. United States, 295 U.S. 495 (1935).

58. *See id.* at 520–21.

59. Act of June 16, 1933, 48 Stat. 195, 196 (quoted in *Schechter Poultry*, 295 U.S. at 521 n.4). The statute required that, before approval of the code or codes, the President find "(1)

a code of fair competition for the live poultry industry. The Code included labor provisions (minimum-wage and maximum-hour requirements) and a trade-practice provision requiring the so-called "straight killing" of poultry.[60]

The government charged the Jewish slaughterhouse operators with various violations of the Code.[61] The slaughterhouse operators responded by challenging—on Article I nondelegation grounds—the underlying statutory grant of authority (from Congress) pursuant to which the President had approved the Code. The challenge succeeded. In declaring the code-making provision of the NIRA to be unconstitutional, the Court explained that "Congress cannot delegate legislative power to the President to exercise an *unfettered discretion* to make whatever laws he thinks may be needed or advisable for the rehabilitation and expansion of trade or industry."[62]

In *Schechter Poultry* and *Panama Refining Co. v. Ryan*,[63] the Court articulated a robust version of the Article I nondelegation doctrine on the way to declaring significant parts of President Roosevelt's New Deal unconstitutional.[64] Unsurprisingly, President Roosevelt was not too pleased with these developments.[65]

---

that such associations or groups impose no inequitable restrictions on admission to membership therein and are truly representative of such trades or industries or subdivisions thereof, and (2) that such code or codes are not designed to promote monopolies or to eliminate or oppress small enterprises and will not operate to discriminate against them, and will tend to effectuate the policy of this title." *Id.*

60. *Schechter Poultry*, 295 U.S. at 525–28. As the Court noted, "[The 'straight killing'] requirement was really one of 'straight' selling. The term 'straight killing' was defined in the code as 'the practice of requiring persons purchasing poultry for resale to accept the run of any half coop, coop, or coops, as purchased by slaughterhouse operators, except for culls.' The charges in the ten counts, respectively, were that the defendants in selling to retail dealers and butchers had permitted 'selections of individual chickens taken from particular coops and half coops.'" *Id.* at 527–28.

61. *Id.* at 525–28.

62. *Id.* at 537–38 (emphasis added).

63. Panama Refining Co. v. Ryan, 293 U.S. 388 (1935). The Court in *Schechter Poultry* cited *Panama Refining*, noting that the case had "recently [provided] occasion to review the pertinent decisions and the general principles which govern the determination of" whether a statute violates the nondelegation doctrine. *Schechter Poultry*, 295 U.S. at 529.

64. *See* Ziaja, *supra* note 38, at 924 (discussing the cases).

65. *See* Franklin D. Roosevelt, U.S. President, Fireside Chat on the Plan for Reorganization of the Judiciary (Mar. 9, 1937), *available at* https://www.presidency.ucsb.edu/documents/fireside-chat-17 [hereinafter FDR Fireside Chat]; David Schoenbrod, *The Delegation Doctrine: Could the Court Give It Substance?*, 83 MICH. L. REV. 1223, 1225 (1985).

After the Court issued its rulings, America voted. In the 1936 election, a referendum on the New Deal, President Roosevelt and the Democrats won in a landslide.[66] Emboldened by the electoral results, President Roosevelt took on the Court. Fed up with the Court's obstruction of his domestic policy agenda, the President proposed a plan by which he would add new Justices to the bench and alter the composition of the Court.[67] President Roosevelt's proposal provided a real-world example of why the stakes for judicial review are so high. William Baude has described this as "the New Deal paradigm," taking the view that "the argument for court reform is especially strong . . . when the Court is standing in the way of Congress, . . . [when] Congress wants to do things, and the [C]ourt won't let them."[68] In one fireside chat in 1937, President Roosevelt opined that "[i]n the last four years the sound rule of giving statutes the benefit of all reasonable [constitutional] doubt has been cast aside. The Court has been acting not as a judicial body, but as a policy-making body."[69] In standing up for the constitutionality of his domestic policies, President Roosevelt launched a political attack on what Alexander Hamilton once described as "the weakest" of the three branches of government.[70]

President Roosevelt's court-packing plan did not come to fruition, perhaps because the Court caved in the face of the political pressure. No surprise—as one writer put it: "The Supreme Court needs a lot of fortitude to challenge one of the elected branches. And in truth, this Court didn't have it."[71] Notably, "[a]fter President Roosevelt threatened to pack the Court if it persisted in rendering such decisions, the Justices changed their tune, and

66. *See 1936: FDR's Second Presidential Campaign*, CUNY: SEE HOW THEY RAN!, http://www.roosevelthouse.hunter.cuny.edu/seehowtheyran/portfolios/1936-fdrs-second-presidential-campaign-the-new-deal (last visited Oct. 29, 2023).

67. *See* FDR Fireside Chat, *supra* note 65.

68. *Settling of Scores*, DIVIDED ARGUMENT, at 16:10 (July 10, 2022), https://www.dividedargument.com/episodes/settling-of-scores.

69. *See* FDR Fireside Chat, *supra* note 65; *see also* Morrison v. Olson, 487 U.S. 654, 724 (1988) (Scalia, J., dissenting) (describing the Court in 1935 as "an activist, anti-New Deal Court bent on reducing the power of President Franklin Roosevelt").

70. THE FEDERALIST NO. 78 (Alexander Hamilton).

71. Peter J. Wallison, *Only the Supreme Court Can Effectively Restrain the Administrative State*, NAT'L REV. (Dec. 1, 2020, 6:30 AM), https://www.nationalreview.com/2020/12/only-the-supreme-court-can-effectively-restrain-the-administrative-state.

nondelegation challenges were thereafter uniformly rejected."[72] In Federalist No. 78, Alexander Hamilton had predicted that such a confrontation would end this way. As Hamilton put it, the judicial branch "can never attack with success either of the other two" branches, and "all possible care is requisite to enable it to defend itself against their attacks."[73]

The Court has not since used the Article I nondelegation doctrine to declare a federal statute unconstitutional. In particular, in the years immediately following 1937, the Court upheld several statutes against nondelegation challenges—including one that was "facially similar to [the] NIRA."[74] By 1944, when the Court decided *Yakus v. United States*,[75] the Article I nondelegation doctrine was effectively a dead letter.[76] To be sure, the nondelegation cases were far from the only reason that President Roosevelt sought to pack the Supreme Court with New Deal sympathizers.[77] But this constitutional episode sheds some light on why the Court became skittish about robust application of the Article I nondelegation doctrine.[78]

72. Merrill, *supra* note 1, at 2103; *see also* Meaghan Dunigan, *The Intelligible Principle: How It Briefly Lived, Why It Died, and Why It Desperately Needs Revival in Today's Administrative State*, 91 ST. JOHN'S L. REV. 247, 259 (2017) ("Both *Panama Refining Co.* and *A.L.A. Schechter Poultry* provide meaningful insight into the intelligible principle. Unfortunately, this insight has largely been dismissed based on the notion that the Court struck down congressional delegation in the [1930s] solely because of the tension that existed between the Court and President Roosevelt. The fact that the Court has not invalidated a statute as an unconstitutional delegation of legislative authority to the executive branch since 1935 largely supports this assertion." (footnote omitted)); George Bunn, Kathleen Irwin & F. Kyra Sido, *No Regulation Without Representation: Would Judicial Enforcement of a Stricter Nondelegation Doctrine Limit Administrative Lawmaking?*, 1983 WIS. L. REV. 341, 342 (pointing to President Roosevelt's court-packing plan as a catalyst for the mid-1930s shift in nondelegation jurisprudence).

73. THE FEDERALIST NO. 78 (Alexander Hamilton).

74. Johnathan Hall, Note, *The Gorsuch Test:* Gundy v. United States*, Limiting the Administrative State, and the Future of Nondelegation*, 70 DUKE L.J. 175, 187–88 (2020).

75. Yakus v. United States, 321 U.S. 414 (1944).

76. *See* Ziaja, *supra* note 38, at 923 ("Scholars furthermore point to *Yakus v. United States* in 1944 as the doctrine's effective end, but the doctrine lost its momentum several years earlier." (footnotes omitted)).

77. *See* Barry Cushman, *The Court-Packing Plan as Symptom, Casualty, and Cause of Gridlock*, 88 NOTRE DAME L. REV. 2089, 2089–90 (2013).

78. Earlier academic literature had predicted the possibility of political backlash to robust judicial review by the federal courts. *See* James B. Thayer, *The Origin and Scope of the American Doctrine of Constitutional Law*, 7 HARV. L. REV. 129, 152 (1893) (suggesting that courts "not step into the shoes of the law-maker").

3. The Decline (and Possible Return?) of the Article I Nondelegation Doctrine: From the Court Packing Plan to the Modern Day

In the years following World War II, the Court sought to articulate a consistent test for disposing of Article I nondelegation challenges. Returning to the pre-1930s regime, the Court recast its nondelegation jurisprudence as being about the application of *J.W. Hampton*'s "intelligible principle" test.[79] But in practice, the "test" was no test at all. In 1974, Justice Marshall put it aptly when he described the Article I nondelegation doctrine as "surely as moribund as the substantive due process approach of the same era—for which the Court is fond of writing an obituary."[80] Some Justices resisted. Perhaps most famously, then-Justice Rehnquist called for the revival of the Article I nondelegation doctrine in a concurrence in what has come to be known as "The Benzene Case."[81] But overall, the nondelegation doctrine failed to gain the support of a majority of the Court.

In 2000, Sunstein suggested that over the years, the Article I nondelegation doctrine has morphed into a set of canons of statutory interpretation. In Sunstein's telling, the doctrine has operated in practice—through the canons—as something of a background constraint on agency action.[82] Given the principle that Congress may not delegate legislative power to the executive branch, Sunstein opines that the nondelegation "canons impose important constraints on administrative authority, for agencies are not permitted to understand ambiguous provisions to give them

79. *See* Mistretta v. United States, 488 U.S. 361, 372 (1989) (describing the "intelligible principle" test as driving the Court's nondelegation jurisprudence in *Panama Refining* and the years following).

80. Fed. Power Comm'n v. New England Power Co., 415 U.S. 352, 353 (1974) (Marshall, J., concurring); *see also* Biden v. Nebraska, 143 S. Ct. 2355, 2378 (2023) (Barrett, J., concurring) ("[T]he 'intelligible principle' test largely leaves Congress to self-police."); *Mistretta*, 488 U.S. at 416 (Scalia, J., dissenting) ("[W]e have almost never felt qualified to second-guess Congress regarding the permissible degree of policy judgment that can be left to those executing or applying the law.").

81. Indus. Union Dep't, AFL-CIO v. Am. Petroleum Inst., 448 U.S. 607, 672, 675 (1980) (Rehnquist, J., concurring).

82. *See* Sunstein, *supra* note 10, at 316.

authority to venture in certain directions; a clear congressional statement is necessary."[83]

One especially important development occurred the next year in 2001, when the Court decided *Whitman v. American Trucking Associations, Inc.*[84] In that case, the Court—unsurprisingly—turned away an Article I nondelegation challenge to a provision of the Clean Air Act.[85] But in so doing, the Court clarified how a nondelegation challenge is supposed to work. Writing for the Court, Justice Scalia explained that "[i]n a delegation challenge, the constitutional question is whether the *statute* has delegated legislative power to the agency."[86] To that end, "an agency [cannot] cure an unlawful delegation of legislative power by adopting in its discretion a limiting construction of the statute."[87] For that reason, litigants lodging Article I nondelegation challenges are challenging the *underlying statute*, not the executive action taken pursuant to the statute's grant of authority.

Fast forward to the present: The Court may be about to revive the Article I nondelegation doctrine. In a 2019 case, *Gundy v. United States*, Justice Gorsuch called for the Court to bring back the doctrine.[88] Chief Justice Roberts and Justice Thomas joined the dissent,[89] and Justice Alito wrote in a concurrence that "[i]f a majority of this Court were willing to reconsider the approach we have taken for the past 84 years, I would support that effort."[90] Now, the Court has two new Justices—Justices Kavanaugh and Barrett—who might be sympathetic to the Article I nondelegation doctrine.[91] That said, not all are convinced that the Court is about

83. *Id.* at 330. In some ways, this observation bears striking similarity to the Court's recent requirement of a clear statement from Congress before an agency can answer a so-called "major question." *See* West Virginia v. EPA, 142 S. Ct. 2587, 2609 (2022) (discussing the "major questions doctrine").

84. Whitman v. Am. Trucking Ass'ns, 531 U.S. 457 (2001).

85. *See id.* at 474.

86. *Id.* at 472 (emphasis added).

87. *Id.*

88. Gundy v. United States, 139 S. Ct. 2116, 2131–33 (2019) (Gorsuch, J., dissenting).

89. *Id.* at 2131.

90. *Id.* (Alito, J., concurring).

91. *See* Peter J. Wallison, *An Empty Attack on the Nondelegation Doctrine*, AM. ENTER. INST. (Apr. 22, 2021), https://www.aei.org/op-eds/an-empty-attack-on-the-nondelegation-doctrine.

to return to declaring statutes unconstitutional on Article I nondelegation grounds.[92]

II. THE BILL OF RIGHTS NONDELEGATION DOCTRINE

Whatever the status of the Article I nondelegation doctrine, the Court has been applying the principles of nondelegation in a related area: Bill of Rights litigation. From the First and Second Amendments to the Fourth and Fifth Amendments, the Court has developed a robust jurisprudence that disfavors the delegation of unfettered discretion. To be sure, not all of the cases or lines of cases discussed in this Part are a *perfect* analog to the traditional understanding of the Article I nondelegation doctrine. But the Bill of Rights nondelegation cases translate the Court's abstract disapproval of delegation into consistently substantive action—in a way that the Article I nondelegation doctrine does not.

This Part reveals a Bill of Rights nondelegation doctrine. It begins by laying out the various cases and lines of cases that come together to form a Bill of Rights nondelegation doctrine. From there, this Part ties the cases together into a coherent doctrinal framework.

A. Laying out the Bill of Rights Nondelegation Doctrine

The Bill of Rights "spells out Americans' rights in relation to their government."[93] These rights are often understood by reference to a particular, substantive guarantee: for example, "Congress shall make no law . . . abridging the freedom of speech";[94] "[t]he right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated";[95] and "private property [shall not] be taken for public use, without just compensation."[96]

92. *See, e.g.*, Thomas A. Koenig & Benjamin R. Pontz, Note, *The Roberts Court's Functionalist Turn in Administrative Law*, 46 HARV. J.L. & PUB. POL'Y 221, 228–30 (2023); Kristin E. Hickman, *The Roberts Court's Structural Incrementalism*, 136 HARV. L. REV. F. 75, 83–84 (2022); Adrian Vermeule, *Never Jam Today*, YALE J. ON REGUL.: NOTICE & COMMENT (June 20, 2019), https://www.yalejreg.com/nc/never-jam-today-by-adrian-vermeule.

93. *The Bill of Rights: What Does It Say?*, NAT'L ARCHIVES, https://www.archives.gov/founding-docs/bill-of-rights/what-does-it-say (last visited Oct. 4, 2023).

94. U.S. CONST. amend. I.

95. U.S. CONST. amend. IV.

96. U.S. CONST. amend. V.

As one scholar has noted, "modern constitutional doctrine has incorporated (almost all of) the Bill of Rights against the states."[97] Today, federal constitutional review of state (and city and municipal) legislation and action is a core component of the Supreme Court's docket.[98] The Fourteenth Amendment's incorporation of the Bill of Rights against the state governments has put the federal courts in the position of rights-guarantor whenever a state abridges the freedom of speech,[99] searches someone's home without a warrant,[100] or effectuates an excessive forfeiture of someone's assets.[101] This doctrine of incorporation recognizes—through the Fourteenth Amendment's Due Process Clause—a set of rigorous individual liberty protections that are enforceable against state governments.

This section—and this Article—builds on the usual understanding of the Bill of Rights. This Article identifies a Bill of Rights "nondelegation doctrine," revealing a coherent framework of cases in which the Supreme Court has constructed a doctrinal edifice of nondelegation around the Bill of Rights. In various cases concerning different Bill of Rights amendments, the Court has evinced hostility to the conferral—or delegation—of too much discretion to the executive to impair certain liberties that the Bill of Rights guarantees. In these cases, the Court has attacked the constitutionality of the underlying *delegation* of discretion to violate an enumerated right (for example, the freedom of speech), as

97. Nicholas Quinn Rosenkranz, *The Objects of the Constitution*, 63 STAN. L. REV. 1005, 1054 (2011).

98. October Term 2020 at the Supreme Court involved various of these sorts of challenges. *See, e.g.*, Americans for Prosperity Found. v. Bonta, 141 S. Ct. 2373 (2021) (First Amendment challenge to state regime of disclosing names of charitable organizations' donors); Cedar Point Nursery v. Hassid, 141 S. Ct. 2063 (2021) (Fifth Amendment Takings Clause challenge to state regulation granting labor organizations a "right to take access" to an agricultural employer's property for union solicitation); Mahanoy Area Sch. Dist. v. B.L., 141 S. Ct. 2038 (2021) (First Amendment challenge to public high school's suspension of student from the cheerleading team because of off-campus speech); Lange v. California, 141 S. Ct. 2011 (2021) (Fourth Amendment challenge to warrantless entry into man's garage after he fled); Fulton v. City of Philadelphia, 141 S. Ct. 1868 (2021) (First Amendment free exercise challenge to city's refusal to contract with Catholic foster care agency unless it agreed to certify same-sex foster couples); Caniglia v. Strom, 141 S. Ct. 1596 (2021) (Fourth Amendment challenge to warrantless removal of a man's firearms from his home); Torres v. Madrid, 141 S. Ct. 989 (2021) (Fourth Amendment challenge to officer shooting a fleeing suspect).

99. *See, e.g.*, Texas v. Johnson, 491 U.S. 397 (1989).

100. *See, e.g.*, Florida v. Jardines, 569 U.S. 1 (2013).

101. *See, e.g.*, Timbs v. Indiana, 139 S. Ct. 682 (2019).

opposed to resting a finding of unconstitutionality solely on the impairment of the liberty itself. In a way, the Bill of Rights nondelegation doctrine operates as a prophylactic, precluding grants of discretion to the executive when that discretion could be used in a way that infringes upon an individual liberty guaranteed by the Bill of Rights.

Doctrinal developments with respect to four separate Bill of Rights amendments illustrate this phenomenon. Beginning with the First Amendment, the "Court has condemned licensing schemes that lodge broad discretion in a public official to permit [or not permit] speech-related activity."[102] Similarly, in a recent Second Amendment decision,[103] the Court declared that a state's discretionary permitting regime for concealed-carry licenses was unconstitutional.[104] The Fourth Amendment also contains a rule of anti-delegation (or at least anti-discretion): various cases have given effect to the Fourth Amendment's requirement that "no [w]arrants shall issue, but upon probable cause, supported by [o]ath or affirmation, and *particularly describing the place to be searched, and the persons or things to be seized*."[105] Moreover, under prevailing Supreme Court precedent, the Fifth Amendment embodies a so-called "void for vagueness doctrine." Here, a court will declare a penal statute to be unconstitutional if it does not "define the criminal offense with sufficient definiteness that ordinary people can understand what conduct is prohibited and in a manner that does not encourage arbitrary and discriminatory enforcement."[106] The vagueness doctrine is therefore "a safeguard against legislative delegation of excessive discretion to courts and

102. Police Dep't of Chi. v. Mosley, 408 U.S. 92, 97 (1972) (first citing Shuttlesworth v. City of Birmingham, 394 U.S. 147 (1969); then citing Cox v. Louisiana, 379 U.S. 536, 555–58 (1965); then citing Staub v. City of Baxley, 355 U.S. 313, 321–25 (1958); and then citing Saia v. New York, 334 U.S. 558, 560–62 (1948)).

103. N.Y. State Rifle & Pistol Ass'n v. Bruen, 142 S. Ct. 2111 (2022).

104. *See id.* at 2122–24.

105. U.S. CONST. amend. IV (emphasis added). For cases applying the principle, see, for example, *Groh v. Ramirez*, 540 U.S. 551, 557 (2004) and *Massachusetts v. Sheppard*, 468 U.S. 981, 988 n.5 (1984).

106. Kolender v. Lawson, 461 U.S. 352, 357 (1983) (first citing Vill. of Hoffman Ests. v. Flipside, Hoffman Ests., Inc., 455 U.S. 489 (1982); then citing Smith v. Goguen, 415 U.S. 566 (1974); then citing Grayned v. City of Rockford, 408 U.S. 104 (1972); then citing Papachristou v. City of Jacksonville, 405 U.S. 156 (1972); and then citing Connally v. Gen. Constr. Co., 269 U.S. 385 (1926)).

to executive officials and agencies, especially the police."[107] Taken together, these cases stand for a broad principle: on the way to enforcing the Bill of Rights, the Court has prevented the delegation of unfettered discretion.

1. Discretionary Licensing Regimes and the First Amendment

In 1948, the Supreme Court considered a First Amendment challenge (via the Fourteenth Amendment) to a local ordinance in Lockport, New York.[108] Samuel Saia was a Jehovah's Witness who wanted "to use sound equipment, mounted atop his car, to amplify lectures on religious subjects."[109] But Lockport law prohibited the use of sound equipment in this way, unless one had obtained permission from the town's chief of police.[110] The local ordinance set no standards for the police chief's issuance of the permit—in other words, issuance of the permit was at the police chief's unfettered discretion.[111] Saia had previously obtained a permit for his use of the sound equipment.[112] But once that permit expired, Saia applied for a new permit, and he was refused.[113] The town grounded its refusal in the fact that some people had apparently complained about Saia.[114] Saia "nevertheless used his equipment as planned on four occasions, but without a permit. He was tried in Police Court for violations of the ordinance."[115]

Saia challenged the constitutionality of the ordinance.[116] By a vote of 5-4 at the Supreme Court, he won.[117] The Court held that the ordinance's permitting regime was unconstitutional, "for it establishe[d] a previous restraint on the right of free speech in violation of the First Amendment which is protected by the

107. Kim Forde-Mazrui, *Ruling out the Rule of Law*, 60 VAND. L. REV. 1497, 1500 (2007) (citing John Calvin Jeffries, Jr., *Legality, Vagueness, and the Construction of Penal Statutes*, 71 VA. L. REV. 189, 215–16 (1985)).

108. Saia v. New York, 334 U.S. 558 (1948).

109. *Id.* at 559.

110. *See id.* at 558 n.1 (citing the local ordinance).

111. *See id.*

112. *See id.* at 559.

113. *Id.*

114. *Id.*

115. *Id.*

116. *Id.* at 558.

117. *See generally id.* (ruling 5-4 in Saia's favor).

Fourteenth Amendment against State action."[118] The Court's main problem with the ordinance was that there were "no standards prescribed for the exercise of [the police chief's] discretion."[119] In the Lockport ordinance, "[t]he right to be heard [was] placed in the uncontrolled discretion of the Chief of Police. He [stood] athwart the channels of communication as an obstruction which [could] be removed only after criminal trial and conviction and lengthy appeal."[120]

The Court analogized the Lockport ordinance to a similar local ordinance that it had declared unconstitutional in *Cantwell v. Connecticut*.[121] The ordinance reviewed in *Cantwell* required that one obtain a license "in order to distribute religious literature."[122] As the Court described the *Cantwell* ordinance in *Saia*: "What was religious was left to the discretion of a public official."[123] The key takeaway from *Saia* was as follows: "When a city allows an official to ban [the use of loud-speakers] in his uncontrolled discretion, it sanctions a device for suppression of free communication of ideas."[124] That suppression—as the Court saw it—is repugnant to the First and Fourteenth Amendments of the Constitution.[125]

Saia was just one of "a series of cases involving discretionary licensing schemes that were, *or might have been*, used to discriminate against certain speech because of its content."[126] The emphasis on "might have been" is important. In *Largent v. Texas*, the Court explained that the very fact that "[d]issemination of ideas depends upon the approval of the distributor by the official . . . . is [itself]

118. *Id.* at 559–60.

119. *Id.* at 560.

120. *Id.* at 560–61.

121. *Id.* at 560 (citing Cantwell v. Connecticut, 310 U.S. 296 (1940)).

122. *Id.* (discussing *Cantwell*). The Free Exercise Clause jurisprudence may also embody a nondelegation principle—in *Fulton v. City of Philadelphia*, Justice Barrett noted in concurrence that "[a] longstanding tenet of [the Court's] free exercise jurisprudence . . . is that a law burdening religious exercise must satisfy strict scrutiny if it *gives* government officials discretion to grant individualized exemptions." 141 S. Ct. 1868, 1883 (2021) (Barrett, J., concurring) (emphasis added) (citing Emp. Div., Dep't of Hum. Res. v. Smith, 494 U.S. 872, 884 (1990)).

123. *Saia*, 334 U.S. at 560.

124. *Id.* at 562.

125. *See id.*

126. Susan H. Williams, *Content Discrimination and the First Amendment*, 139 U. PA. L. REV. 615, 627 n.42 (1991) (emphasis added); *see also* Police Dep't of Chi. v. Mosley, 408 U.S. 92, 97 (1972) (collecting cases).

administrative censorship in an extreme form" that violates the Constitution.[127] The point of the First Amendment discretionary licensing cases is that the "lodg[ing of] broad discretion in a public official to permit speech-related activity" itself abridges speech.[128] The constitutional problem is the licensing schemes' "potential use as instruments for selectively suppressing some points of view."[129] That mere potential is enough to create a constitutional difficulty.

After *Saia*,[130] the Supreme Court confronted a number of other discretionary licensing regimes that imperiled First (and Fourteenth) Amendment rights. In a pair of cases decided on the same day in early 1951—*Niemotko v. Maryland*[131] and *Kunz v. New York*[132]—the Court expanded upon its First Amendment discretionary permitting jurisprudence.[133] The ordinances at issue in the cases had vested an unfettered discretion in local officials to deny permits, which the Court found unacceptable.[134] In the words of the *Kunz* Court, a state "cannot vest restraining control over the right to speak . . . in an administrative official where there are no

127. Largent v. Texas, 318 U.S. 418, 422 (1943).

128. *Mosley*, 408 U.S. at 97 (first citing Shuttlesworth v. City of Birmingham, 394 U.S. 147 (1969); then citing Cox v. Louisiana, 379 U.S. 536, 555–58 (1965); then citing Staub v. City of Baxley, 355 U.S. 313, 321–25 (1958); and then citing *Saia*, 334 U.S. at 560–62).

129. *Id.* At times, the Supreme Court has recognized the constitutional difficulty and instead interpreted federal law narrowly to avoid the conclusion that Congress granted unfettered discretion to an executive official when the First Amendment was on the line. In *Kent v. Dulles*, for example, the Court confronted the U.S. Secretary of State's denials of passports to suspected communists. *See* Kent v. Dulles, 357 U.S. 116, 117–19 (1958). Instead of declaring the underlying passport-granting laws unconstitutional, the Court avoided the constitutional question and determined that "[i]t would . . . be strange to infer that . . . the Secretary has been silently granted by Congress the . . . power to curtail in his discretion the free movement of citizens in order to satisfy himself about their beliefs or associations." *Id.* at 130.

130. To be sure, *Saia* was not the first case to condemn discretionary licensing regimes in the speech context. *See, e.g.*, Thornhill v. Alabama, 310 U.S. 88, 97 (1940); Schneider v. State, 308 U.S. 147, 164 (1939); Hague v. Comm. for Indus. Org., 307 U.S. 496, 516 (1939); *see also* Lovell v. City of Griffin, 303 U.S. 444, 451–52 (1938) (describing a discretionary licensing scheme as inconsistent with the freedom of the press guaranteed by the First Amendment).

131. Niemotko v. Maryland, 340 U.S. 268 (1951).

132. Kunz v. New York, 340 U.S. 290 (1951).

133. *See Niemotko*, 340 U.S. at 273; *Kunz*, 340 U.S. at 295.

134. *See Niemotko*, 340 U.S. at 273; *Kunz*, 340 U.S. at 295. The Court noted in *Kunz* that it had "consistently condemned licensing systems which vest in an administrative official discretion to grant or withhold a permit upon broad criteria unrelated to proper regulation of public places." *Kunz*, 340 U.S. at 294.

appropriate standards to guide his action."[135] Over the next few decades, the Court continued to enforce this principle.[136] In case after case, regardless of whether the laws were neutral with respect to the expressive activity's message, "the Court was worried about the broad discretion they gave to government officials. The Court's suspicion of such discretion arose, in large part, from its fear that officials would use their power to discriminate among speakers based upon the content of their speech."[137]

Instead of making litigants challenge these discretion-delegating schemes in an as-applied posture, the Court has permitted facial constitutional attacks.[138] By 1988, the Court explained that its "cases have long held that when a licensing statute allegedly vests unbridled discretion in a government official over whether to permit or deny expressive activity, one who is subject to the law may challenge it facially without the necessity of first applying for, and being denied, a license."[139] This constitutional approach flows from "the time-tested knowledge

135. *Kunz*, 340 U.S. at 295.

136. *See, e.g.*, Sec'y of Md. v. Joseph H. Munson Co., 467 U.S. 947, 964 n.12 (1984); Shuttlesworth v. City of Birmingham, 394 U.S. 147, 150–51 (1969); Freedman v. Maryland, 380 U.S. 51, 56–57 (1965); Cox v. Louisiana, 379 U.S. 536, 556–58 (1965); Staub v. City of Baxley, 355 U.S. 313, 325 (1958).

137. Williams, *supra* note 126, at 701.

138. Richard Fallon explains the difference between "as-applied" and "facial" challenges:

> In an as-applied challenge, a party maintains that the Constitution forbids a statute's application to his or her case. In contrast, a facial challenge asserts that a statute—or, more commonly, a provision of a multipart statute—exhibits a defect that renders it invalid as applied to all cases, even if a more narrowly (or occasionally a more broadly) framed provision could have prohibited the challenger's conduct.

Richard H. Fallon, Jr., *Facial Challenges, Saving Constructions, and Statutory Severability*, 99 TEX. L. REV. 215, 228 (2020) (footnote omitted).

139. City of Lakewood v. Plain Dealer Pub. Co., 486 U.S. 750, 755–56 (1988) (first citing *Freedman*, 380 U.S. at 56; then citing Thornhill v. Alabama, 310 U.S. 88, 97 (1940); then citing *Shuttlesworth*, 394 U.S. at 151; then citing Lovell v. City of Griffin, 303 U.S. 444, 452–53 (1938); and then citing *Joseph H. Munson Co.*, 467 U.S. at 956–57); *see also* FW/PBS, Inc. v. City of Dallas, 493 U.S. 215, 246 (1990) (White, J., concurring in part and dissenting in part) ("Licensing schemes subject to First Amendment scrutiny . . . have been invalidated when undue discretion has been vested in the licensor. Unbridled discretion with respect to the criteria used in deciding whether or not to grant a license is deemed to convert an otherwise valid law into an unconstitutional prior restraint. That rule reflects settled law with respect to licensing in the First Amendment context." (first citing *Shuttlesworth*, 394 U.S. at 150–52; then citing *Lakewood*, 486 U.S. at 757; then citing *Staub*, 355 U.S. 313; then citing *Niemotko*, 340 U.S. 268; then citing *Kunz*, 340 U.S. 290; and then citing Saia v. New York, 334 U.S. 558 (1948))).

that in the area of free expression a licensing statute placing unbridled discretion in the hands of a government official or agency constitutes a prior restraint and may result in censorship."[140] In the Court's telling, it is "the mere existence of the licensor's unfettered discretion, coupled with the power of prior restraint, [that] intimidates parties into censoring their own speech, even if the discretion and power are never actually abused."[141]

From these cases, a rule of nondelegation for the First Amendment emerges: A statute is unconstitutional if it delegates standardless discretion to a government official to permit or deny expressive activity.

2. "May Issue" Concealed Carry Permitting Regimes and the Second Amendment

The Court recently recognized something of a rule of nondelegation in a Second Amendment case, too. While this rule does not have anything close to the doctrinal pedigree of the First Amendment nondelegation principle discussed earlier in this section, it nevertheless contributes to this Article's identification of a Bill of Rights nondelegation doctrine.

In October Term 2021, the Supreme Court decided *New York State Rifle & Pistol Association v. Bruen*.[142] The case concerned a New York state licensing regime for concealed-carry permits. To set the stage, New York law prohibited possession of a firearm without a

140. *Lakewood*, 486 U.S. at 757 (first citing *Shuttlesworth*, 394 U.S. at 151; then citing *Cox*, 379 U.S. 536; then citing *Staub*, 355 U.S. at 321–22; then citing *Kunz*, 340 U.S. at 294; then citing *Niemotko*, 340 U.S. 268; and then citing *Saia*, 334 U.S. 558).

141. *Id.* Perhaps in these sorts of cases, the Court is worried about what it describes as "the more covert forms of discrimination that may result when arbitrary discretion is vested in some governmental authority." Heffron v. Int'l Soc'y for Krishna Consciousness, Inc., 452 U.S. 640, 649 (1981). Susan Williams takes it a step further, arguing that "[t]he serious, present harm in . . . discretion lies . . . in the concept of chill, a concept more closely related to content discrimination from the speaker's perspective." Williams, *supra* note 126, at 704; *see also Joseph H. Munson Co.*, 467 U.S. at 964 n.12 ("By placing discretion in the hands of an official to grant or deny a license, such a statute creates a threat of censorship that by its very existence chills free speech."(first citing *Thornhill*, 310 U.S. at 97; then citing Schneider v. New Jersey, 308 U.S. 147 (1939); and then citing *Lovell*, 303 U.S. at 451)). As Williams puts it, "Curing or avoiding chill requires changing the regulatory scheme so that speakers no longer feel threatened rather than changing the motives of government actors, who, by hypothesis, need not actually be discriminating." Williams, *supra* note 126, at 704.

142. N.Y. State Rifle & Pistol Ass'n v. Bruen, 142 S. Ct. 2111 (2022).

license.[143] To obtain a license for carrying a firearm outside the home, an applicant had to prove that "proper cause" existed for the license's issuance.[144] As the Court noted, "[n]o New York statute defines 'proper cause,'" although New York courts had interpreted the term to mean "a special need for self-protection distinguishable from that of the general community."[145] Unfortunately for applicants, judicial review of these licensing decisions was limited.[146] New Yorkers Brandon Koch and Robert Nash had applied for unrestricted licenses to carry firearms, but licensing officials only issued them restricted permits.[147]

Koch and Nash sued the relevant state officials, alleging violations of their Second and Fourteenth Amendment rights.[148] Like Samuel Saia and many other litigants in the First Amendment cases discussed in section II.A.1, Koch and Nash won. The Court noted that it had "granted certiorari to decide whether New York's *denial* of [Koch and Nash's] license applications violated the Constitution."[149] But the Court ultimately determined that it was the state's "proper-cause requirement" that violated the Fourteenth Amendment's incorporation of the Second Amendment liberty.[150] The Court compared New York's licensing scheme—"under which authorities have *discretion* to deny concealed-carry licenses even when the applicant satisfies the statutory criteria, usually because the applicant has not demonstrated cause or suitability for the relevant license[]"—to "the vast majority of States . . . where authorities must issue concealed-carry licenses whenever applicants satisfy certain threshold requirements, without granting licensing officials *discretion* to deny licenses based on a perceived lack of need or suitability."[151] The contrast here was between "may-issue" and "shall-issue" regimes; in "may-issue" regimes, the legislatures give discretion to licensing officials to deny permits even if applicants meet the statutory criteria.

143. *See id.* at 2122.

144. *See id.* at 2123 (citing N.Y. PENAL LAW § 400.00(2)(f) (McKinney 2006)).

145. *Id.* (quoting Klenosky v. N.Y.C. Police Dep't, 428 N.Y.S.2d 256, 257 (N.Y. App. Div. 1980)).

146. *See id.*

147. *Id.* at 2125.

148. *Id.*

149. *Id.* (emphasis added).

150. *Id.* at 2156.

151. *Id.* at 2123–24 (emphasis added).

Justice Kavanaugh's concurrence specifically addressed the discretionary nature of "may-issue" licensing regimes.[152] He noted that "[a]s the Court explain[ed], New York's outlier may-issue regime is constitutionally problematic because it grants open-ended discretion to licensing officials and authorizes licenses only for those applicants who can show some special need apart from self-defense."[153] Echoing the themes of the First Amendment licensing cases, Justice Kavanaugh wrote that "the unchanneled discretion for licensing officials and the special-need requirement . . . in effect deny the right to carry handguns for self-defense to many 'ordinary, law-abiding citizens.'"[154] Here again, the very delegation of discretion constituted a denial of the enumerated right.

The *Bruen* Court also took issue with New York State's declaration of the island of Manhattan as a "sensitive place" "where the government may lawfully disarm law-abiding citizens" consistent with the Second Amendment.[155] Further, the Court concluded that the respondents in the case had "failed to meet their burden to identify an American tradition justifying New York's proper-cause requirement" before deeming the requirement unconstitutional.[156] Nevertheless, the discussion of discretion in *Bruen* provides further evidence for the existence of a Bill of Rights nondelegation doctrine.

Bruen—or, at least, Justice Kavanaugh's concurrence in *Bruen*—continued the Court's march of anti-delegation through the Bill of Rights. Like the First Amendment cases, *Bruen* stands for a nondelegation principle: legislatures may not delegate open-ended discretion to (executive) licensing officials to deny concealed-carry permits.

3. General Warrants and the Fourth Amendment

The Fourth Amendment spells out its anti-delegation rule more explicitly than do the First and Second Amendments. After articulating an overarching prohibition on unreasonable government searches and seizures, the Fourth Amendment provides that "no

152. *See id.* at 2161 (Kavanaugh, J., concurring).

153. *Id.*

154. *Id.* (quoting *id.* at 2122 (majority opinion)) (citing District of Columbia v. Heller, 554 U.S. 570, 635 (2008)).

155. *Id.* at 2133–34 (majority opinion).

156. *Id.* at 2138.

[w]arrants shall issue, but upon probable cause, supported by [o]ath or affirmation, and *particularly describing the place to be searched, and the persons or things to be seized*."[157] The Fourth Amendment's particularity requirement sets up a two-part test: the court must ask whether a warrant particularly describes (1) the *place* to be *searched* and (2) the *persons or things* to be *seized*. If the answer to either question is "no," the warrant is "invalid."[158] And "[a]lthough the Fourth Amendment does not, by its text, require that searches be supported by a warrant, '[the Supreme] Court has inferred that a [valid] warrant must generally be secured' for a search to comply with the Fourth Amendment."[159]

Remedies are tricky in Fourth Amendment cases. The ordinary remedy for a Fourth Amendment violation is the application of the so-called "exclusionary rule," under which evidence obtained in violation of one's Fourth Amendment rights is inadmissible against that person at a criminal trial.[160] But in recognition of "the substantial [social] costs" of excluding evidence, the Supreme Court has carved out numerous exceptions to the exclusionary rule.[161]

One of these carve-outs is known as the "good-faith exception," in which a court will not apply the exclusionary rule when an officer acts in "objectively reasonable reliance" on an invalid warrant.[162] And while the Court has applied the good-faith exception

157. U.S. CONST. amend. IV (emphasis added).

158. Groh v. Ramirez, 540 U.S. 551, 557 (2004).

159. Mitchell v. Wisconsin, 139 S. Ct. 2525, 2539–40 (2019) (Thomas, J., concurring in the judgment) (quoting Kentucky v. King, 563 U.S. 452, 459 (2011)) (citing *Groh*, 540 U.S. at 571–73 (Thomas, J., dissenting)); *see also Groh*, 540 U.S. at 559 ("[T]he presumptive rule against warrantless searches applies with equal force to searches whose only defect is a lack of particularity in the warrant.").

160. *See* Edwin G. Fee, Jr., *Criminal Procedure I: Narrowing the Protection of the Fourth Amendment*, 1989 ANN. SURV. AM. L. 371, 371 & n.4; *see also* Alexandra Natapoff, Atwater *and the Misdemeanor Carceral State*, 133 HARV. L. REV. F. 147, 166 (2020) ("The usual rule is that police cannot use the fruits of an illegal seizure." (citing Utah v. Strieff, 136 S. Ct. 2056, 2061 (2016))). While the Court in 1971 recognized a private right of action for money damages under the Fourth Amendment itself, *see* Bivens v. Six Unknown Named Agents of the Fed. Bureau of Narcotics, 403 U.S. 388, 397–98 (1971), the Court has pared that remedy back in recent years. *See, e.g.*, Vega v. Tekoh, 142 S. Ct. 2095 (2022); Hernandez v. Mesa, 140 S. Ct. 735 (2020).

161. United States v. Leon, 468 U.S. 897, 922–23 (1984); *see also* Hudson v. Michigan, 547 U.S. 586, 591–92 (2006) (describing a balancing approach for application of the exclusionary rule).

162. *Leon*, 468 U.S. at 927 (Blackmun, J., concurring).

notwithstanding a particularity violation,[163] the exclusionary rule still undoubtedly applies when "a warrant [is] so facially deficient—*i.e.*, in failing to particularize the place to be searched or the things to be seized—that the executing officers cannot reasonably presume it to be valid."[164] Thus, in light of the Court's exception-laden exclusionary rule jurisprudence, the particularity requirement remains an important bulwark against judicial delegation of discretion to police.[165] And when a warrant is insufficiently particular, the judge's issuance of the warrant permits the police to fill up the details, on their own, as they see fit—a discretionary exercise of the police power that the Fourth Amendment disallows.

The Supreme Court has been clear about this principle: "The uniformly applied rule is that a search conducted pursuant to a warrant that fails to conform to the particularity requirement of the Fourth Amendment is unconstitutional."[166] Conformance to the particularity requirement requires that "nothing is left to the discretion of the officer executing the warrant."[167] Particularity in the application for the warrant is not enough.[168] The warrant itself must provide the requisite particularity. To be sure, the "particularity requirement does not include the conditions precedent to execution of the warrant"—that is, anticipatory warrants (warrants with a

163. *See* Massachusetts v. Sheppard, 468 U.S. 981 (1984); *see also generally* Martha Applebaum, Note, *"Wrong but Reasonable": The Fourth Amendment Particularity Requirement After* United States v. Leon, 16 FORDHAM URB. L.J. 577 (1987) (analyzing application of the good-faith exception to the exclusionary rule in cases concerning particularity-deficient warrants in the years immediately following *Sheppard*).

164. *Leon*, 468 U.S. at 923 (majority opinion) (citing *Sheppard*, 468 U.S. at 988–91).

165. *But cf.* David Alan Sklansky, *The Nature and Function of Prosecutorial Power*, 106 J. CRIM. L. & CRIMINOLOGY 473, 505 (2016) ("Police discretion is hemmed in only at the margins by legal constraints . . . ." (first citing Egon Bittner, *The Police on Skid-Row: A Study of Peace Keeping*, 5 AM. SOC. REV. 699 (1967); then citing Ryken Grattet & Valerie Jenness, *The Reconstitution of Law in Local Settings: Agency Discretion, Ambiguity, and a Surplus of Law in the Policing of Hate Crime*, 39 LAW & SOC'Y REV. 893, 899–900 (2005); and then citing David A. Sklansky, *Killer Seatbelts and Criminal Procedure*, 119 HARV. L. REV. F. 56, 58–59 (2006)).

166. *Sheppard*, 468 U.S. at 988 n.5 (first citing Stanford v. Texas, 379 U.S. 476 (1965); then citing United States v. Cardwell, 680 F.2d 75, 77–78 (9th Cir. 1982); then citing United States v. Crozier, 674 F.2d 1293, 1299 (9th Cir. 1982); then citing United States v. Klein, 565 F. 2d 183, 185 (1st Cir. 1977); then citing United States v. Gardner, 537 F.2d 861, 862 (6th Cir. 1976); and then citing United States v. Marti, 421 F.2d 1263, 1268–69 (2d. Cir. 1970)).

167. Marron v. United States, 275 U.S. 192, 196 (1927).

168. *See* Groh v. Ramirez, 540 U.S. 551, 557 (2004).

"triggering condition") are constitutionally valid, even if the warrant does not itself specify the triggering condition.[169]

The particularity requirement is rooted in the Framers' abhorrence of the "general warrant."[170] Also known as "writs of assistance," general warrants in the colonies gave British "customs officials blanket authority to search where they pleased for goods imported in violation of the British tax laws."[171] General warrants would authorize standardless searches and seizures, drawing the ire of the colonists.[172] Moreover, officials used general warrants to harass dissenters.[173] For this reason, the particularity requirement has a doctrinal connection to the First Amendment's protections for speech and expression.[174] In *Stanford v. Texas*, the Court expounded on this connection when it explained that "the constitutional requirement that warrants must particularly describe the 'things to be seized' is to be accorded the most scrupulous exactitude when the 'things' are books, and the basis for their seizure is the ideas which they contain."[175] In fact, resistance against the practice of general warrants may have been the spark that ignited the revolutionary fire in colonial America.[176]

The Fourth Amendment's particularity requirement is a rule of nondelegation. Unlike the usual delegation scenario, the delegation in the Fourth Amendment particularity requirement is from a

169. United States v. Grubbs, 547 U.S. 90, 98 (2006). *But cf. The Supreme Court – Leading Cases*, 120 HARV. L. REV. 125, 161 (2006) (describing anticipatory warrants as constituting an "inherent delegation of discretion from an impartial magistrate to the officer executing the warrant"). Indeed, this carveout is similar to the Article I nondelegation doctrine's understanding that "once Congress prescribes the rule governing private conduct, it may make the application of that rule depend on executive fact-finding." Gundy v. United States, 139 S. Ct. 2116, 2136 (2019) (Gorsuch, J., dissenting).

170. Coolidge v. New Hampshire, 403 U.S. 443, 467 (1971) (first citing Boyd v. United States, 116 U.S. 616, 624–30 (1886); then citing *Marron*, 275 U.S. at 195–96; and then citing *Stanford*, 379 U.S. at 476).

171. *Stanford*, 379 U.S. at 481.

172. *See id.*

173. *See* Marcus v. Search Warrant of Property, 367 U.S. 717, 724–25 (1961).

174. *See id.* at 729 ("The Bill of Rights was fashioned against the background of knowledge that unrestricted power of search and seizure could also be an instrument for stifling liberty of expression. For the serious hazard of suppression of innocent expression inhered in the discretion confided in the officers authorized to exercise the power.").

175. *Stanford*, 379 U.S. at 485–86 (first citing *Marcus*, 367 U.S.; and then citing A Quantity of Books v. Kansas, 378 U.S. 205 (1964)).

176. Boyd v. United States, 116 U.S. 616, 625 (1886).

judicial officer to an executive officer.[177] True, one might conceive of the requirement of the warrant itself as a limitation on executive power.[178] But whether the discretionary constitutional authority to issue the warrant is a freestanding element of the judicial power or a cabining of the executive power, the particularity requirement ensures that the judicial branch does not merely delegate the discretionary authority (to determine whether the warrant is particular enough) *back* to the executive.

The principle of nondelegation therefore holds up. A main problem with general warrants was that "they delegated to the officer the power to decide whom to search and for what to search."[179] At bottom, general warrants constituted a "delegation of discretion."[180] To rectify this issue, the Framers enshrined a requirement in the Fourth Amendment that warrants "particularly describ[e] the place to be searched, and the persons or things to be seized."[181] The Court has viewed "[t]he security of one's privacy against arbitrary intrusion by the police" as being "at the core of the

177. *Cf.* Johnson v. United States, 333 U.S. 10, 14 (1948) ("When the right of privacy must reasonably yield to the right of search is, as a rule, to be decided by a judicial officer, not by a policeman or government enforcement agent."). Courts have recognized the possibility of judicial delegations to the executive in other contexts. *See, e.g.*, United States v. Matta, 777 F.3d 116, 121 (2d Cir. 2015) (upholding a defendant's challenge to his sentence when "the District Court impermissibly delegated its sentencing authority by allowing the Probation Department to determine whether he should undergo inpatient or outpatient drug treatment as a condition of supervised release"). The delegation can go the other way, too. *See, e.g.*, Dave Yost & Benjamin M. Flowers, *The First Step Act and the Pardon Power*, HARV. J.L. & PUB. POL'Y PER CURIAM (No. 42, Fall 2023), at 2 (contending that the First Step Act's compassionate-release provisions unconstitutionally delegated the President's pardon power to the judiciary).

178. *See, e.g.*, Lo-Ji Sales, Inc. v. New York, 442 U.S. 319, 326 (1979) (describing "a warrant authorized by a neutral and detached judicial officer" as "a more reliable safeguard against improper searches than the hurried judgment of a law enforcement officer" (quoting United States v. Chadwick, 433 U.S. 1, 9 (1977)) (citing Coolidge v. New Hampshire, 403 U.S. 443, 450 (1971))).

179. Barbara C. Salken, *The General Warrant of the Twentieth Century? A Fourth Amendment Solution to Unchecked Discretion to Arrest for Traffic Offenses*, 62 TEMP. L. REV. 221, 254 (1989).

180. *Id.*; *cf.* City of Los Angeles v. Patel, 576 U.S. 409, 427 (2015) (describing police discretion as an evil against which the Fourth Amendment guards); Anthony G. Amsterdam, *Perspectives on the Fourth Amendment*, 58 MINN. L. REV. 349, 412 (1974) ("Under the fourth amendment, even where the initial justification for a search was determined by a magistrate, executive discretion in its execution was to be curbed by the requirement of particularity of description in the warrant of the items subject to seizure.").

181. U.S. CONST. amend. IV.

Fourth Amendment" and "basic to a free society."[182] General warrants countenance arbitrary police intrusion, eroding "[t]he right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures"[183] because the police become the ones who get to fill out the substance of the warrant in practice through their searching. These warrants violate the Bill of Rights nondelegation doctrine.

4. The Void-for-Vagueness Doctrine and the Fifth Amendment

The Bill of Rights doctrine most familiar to the nondelegation discourse is the so-called "void-for-vagueness" doctrine. Various scholars and jurists have linked the void-for-vagueness doctrine to the Article I nondelegation doctrine.[184] The two have some overlap, to be sure. But the better way to understand the vagueness doctrine is that it exists as a component part of a wholly different nondelegation doctrine: the Bill of Rights nondelegation doctrine.

The Fifth Amendment provides, in relevant part, that "[n]o person shall . . . be deprived of life, liberty, or property, without due process of law . . . ."[185] Similarly, the Fourteenth Amendment prohibits "any [s]tate" from "depriv[ing] any person of life, liberty, or property, without due process of law[.]"[186] In these clauses, the Supreme Court has found a rule of constitutional law applicable against the federal government via the Fifth Amendment and analogously applicable to the state governments via the Fourteenth Amendment. That rule is as follows: "[T]he Government violates [the Constitution's guarantee of due process] by taking away someone's life, liberty, or property under a criminal law so vague that it fails to give ordinary people fair notice of the conduct it punishes, or so standardless that it invites arbitrary enforcement."[187] The Court recently even applied the void-for-vagueness doctrine to

182. Wolf v. Colorado, 338 U.S. 25, 27 (1949).

183. U.S. CONST. amend. IV.

184. *See, e.g.*, *supra* notes 19–20; *infra* notes 202–205; Douglas H. Ginsburg & Steven Menashi, *Nondelegation and the Unitary Executive*, 12 U. PA. J. CONST. L. 251, 264 n.72 (2010).

185. U.S. CONST. amend. V.

186. U.S. CONST. amend. XIV.

187. Johnson v. United States, 576 U.S. 591, 595 (2015) (applying the Fifth Amendment's Due Process Clause in evaluating the constitutionality of a federal law); Kolender v. Lawson, 461 U.S. 352, 357–58 (1983) (applying the Fourteenth Amendment's Due Process Clause in evaluating the constitutionality of a state law).

a civil law when the consequence of the civil penalty in question was deportation.[188] At bottom, "the 'void for vagueness' doctrine requires the state to set forth clear guidance before it may punish private conduct."[189]

Like the Fourth Amendment's particularity requirement, the void-for-vagueness doctrine has a special connection to the First Amendment. Despite the traditional understanding that the vagueness doctrine is reserved for criminal laws, Justice Thomas noted in concurrence in *Johnson v. United States* that the Court had previously applied the vagueness doctrine to a non-penal law.[190] The case he cited for this proposition was *Keyishian v. Board of Regents*,[191] in which the Court held that multiple state laws were unconstitutional on vagueness grounds. But the decision was quite focused on the freedoms that the First Amendment guarantees. The state laws at issue would have authorized the removal of state education employees who uttered treasonous or seditious words.[192] In making its declaration of unconstitutionality, the *Keyishian* Court cited *NAACP v. Button*, in which the Court had explained that "standards of permissible statutory vagueness are strict in the area of free expression."[193] Moreover, in *Grayned v. City of Rockford*,[194] the Court drew a connection between its vagueness holdings and its First Amendment licensing cases (discussed in section II.A.1).[195] As Sunstein has written, specificity in the state's prescriptions of rules

188. *See* Sessions v. Dimaya, 138 S. Ct. 1204 (2018).

189. Cass R. Sunstein, *Problems with Rules*, 83 CALIF. L. REV. 953, 968 (1995) [hereinafter Sunstein, *Rules*] (first citing Papachristou v. City of Jacksonville, 405 U.S. 156, 162 (1972); then citing Gooding v. Wilson, 405 U.S. 518, 520–21 (1972); then citing Baggett v. Bullitt, 377 U.S. 360, 367–70 (1964); and then citing Smith v. Goguen, 415 U.S. 566, 572–73 (1974)).

190. *Johnson*, 576 U.S. at 612 (Thomas, J., concurring) (citing Keyishian v. Bd. of Regents, 385 U.S. 589, 597–604 (1967)).

191. Keyishian v. Bd. of Regents, 385 U.S. 589 (1967).

192. *See id.* at 593.

193. NAACP v. Button, 371 U.S. 415, 432 (1963) (first citing Smith v. California, 361 U.S. 147, 151 (1959); then citing Winters v. New York, 333 U.S. 507, 509–10, 517–18 (1948); then citing Herndon v. Lowry, 301 U.S. 242 (1937); then citing Stromberg v. California, 283 U.S. 359 (1931); and then citing United States v. CIO, 335 U.S. 106, 142 (1948) (Rutledge, J., concurring in the result)) (quoted in *Keyishian*, 385 U.S. at 604).

194. Grayned v. City of Rockford, 408 U.S. 104 (1972).

195. *See id.* at 113 n.22.

might be seen as "particularly important in the areas of criminal justice and freedom of speech[.]"[196]

The justifications for the vagueness doctrine are twofold. First, the vagueness doctrine guarantees that people will have "fair notice" of what conduct is proscribed.[197] Second, and pertinent to this Article's thesis, "the doctrine guards against arbitrary or discriminatory law enforcement by insisting that a statute provide standards to govern the actions of police officers, prosecutors, juries, and judges."[198] This second justification—which itself informs the first justification—is undergirded by a kind of nondelegation rationale. Here, the Court's concern has been that "[s]tatutory language of . . . a standardless sweep allows policemen, prosecutors, and juries to pursue their personal predilections. Legislatures *may not so abdicate their responsibilities* for setting the standards of the criminal law."[199] As Justice Kagan observed in a recent plurality opinion, "[T]he doctrine is a corollary of the separation of powers—requiring that Congress, rather than the executive or judicial branch, define what conduct is

196. Sunstein, *Rules*, *supra* note 189, at 968. F. Andrew Hessick and Carissa Byrne Hessick would go a step further. They argue that there is an "incompatibility between the prevailing justification for modern nondelegation doctrine and the vagueness doctrine . . . ." F. Andrew Hessick & Carissa Byrne Hessick, *Nondelegation and Criminal Law*, 107 VA. L. REV. 281, 286 (2021). In making this point, they posit that "treating criminal delegations no differently than other delegations" is a "fundamental problem"—as they put it, "criminal law delegations are different from other delegations. They are inconsistent with foundational criminal law doctrine, they present greater threats to the principles underlying the nondelegation doctrine, and they are not supported by the ordinary arguments in favor of delegation. And so we should treat criminal law delegations differently." *Id.* (emphasis omitted).

Justice Gorsuch has linked the Fourth Amendment's particularity requirement to the vagueness doctrine. *See* Sessions v. Dimaya, 138 S. Ct. 1204, 1227 (2018) (Gorsuch, J., concurring). Some scholars have also drawn connections between the Fourth Amendment and the vagueness doctrine. *See, e.g.*, Forde-Mazrui, *supra* note 107, at 1500 n.27; Tracey Maclin, *What Can Fourth Amendment Doctrine Learn from Vagueness Doctrine?*, 3 U. PA. J. CONST. L. 398, 404 (2001).

197. *See Dimaya*, 138 S. Ct. at 1212 (plurality opinion).

198. *Id.* (citing Kolender v. Lawson, 461 U.S. 352, 357–58 (1983)); *see also* Smith v. Goguen, 415 U.S. 566, 574 (1974) ("[P]erhaps the most meaningful aspect of the vagueness doctrine is not actual notice, but the other principal element of the doctrine—the requirement that a legislature establish minimal guidelines to govern law enforcement.").

199. *Smith*, 415 U.S. at 575 (emphasis added) (citing Papachristou v. City of Jacksonville, 405 U.S. 156, 165–69 (1972)); *see also Papachristou*, 405 U.S. at 168 ("Another aspect of the ordinance's vagueness appears when we focus, not on the lack of notice given a potential offender, but on the effect of the unfettered discretion it places in the hands of the . . . police.").

sanctionable and what is not."[200] And when Congress (or a state legislature) has failed to do so, thereby delegating this awesome power to the executive and the courts under a broad grant of penal authority, the Court has not hesitated to declare the offending statutes unconstitutional.[201]

The Article I nondelegation parallels are evident. Dissenting in *Gundy v. United States*, Justice Gorsuch submitted that the Court sometimes uses the vagueness doctrine in place of the nondelegation doctrine to "rein in Congress's efforts to delegate legislative power."[202] And dissenting in *Sessions v. Dimaya*, Justice Thomas hypothesized that "the vagueness doctrine is really a way to enforce the separation of powers—specifically, the doctrine of nondelegation."[203] To make this point, he pointed to the Court's admonition in *Grayned* that "[a] vague law impermissibly delegates basic policy matters."[204] Justice Thomas noted that he locates the nondelegation principle, which he defined as the rule "that the Constitution prohibits Congress from delegating core legislative power to another branch[,]" "in the Vesting Clauses of Articles I, II, and III—not in the Due Process Clause."[205] This view of the meaning of the nondelegation principle comports with the present scholarly discourse. Yet as this Article demonstrates, the rule that Justice Thomas describes is merely one *kind* of nondelegation

200. *Dimaya*, 138 S. Ct. at 1212; *see also* City of Chicago v. Morales, 527 U.S. 41, 70 (1999) (Breyer, J., concurring) ("[I]t is in the ordinance's delegation to the policeman of open-ended discretion . . . that the problem lies.").

201. *See* Johnson v. United States, 576 U.S. 591, 612 (2015) (Thomas, J., concurring) ("We have become accustomed to using the Due Process Clauses to invalidate laws on the ground of 'vagueness.' The doctrine we have developed is quite sweeping . . . . Using this framework, we have nullified a wide range of enactments." (citation omitted)) (collecting cases).

202. Gundy v. United States, 139 S. Ct. 2116, 2141–42 (2019) (Gorsuch, J., dissenting).

203. *Dimaya*, 138 S. Ct. at 1248 (Thomas, J., dissenting). The idea that other aspects of law have "replaced" the nondelegation doctrine is not limited to vagueness. *See, e.g.*, Nathan Richardson, *Antideference: COVID, Climate, and the Rise of the Major Questions Canon*, 108 VA. L. REV. ONLINE 174, 206 (2022) ("Instead of avoiding the difficulties of applying the nondelegation doctrine, the major questions canon achieves the same purpose *sub rosa*."); Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249, 1295 (2008) ("[T]he [Administrative Procedure Act's] procedural constraints on the exercise of delegated discretion have effectively replaced the nondelegation doctrine."); *cf.* Dan M. Kahan, *Lenity and Federal Common Law Crimes*, 1994 SUP. CT. REV. 345, 345 ("Narrow construction of criminal statutes, it is proclaimed, . . . constrains the discretion of law enforcement officials . . . .").

204. *Dimaya*, 138 S. Ct. at 1248 (Thomas, J., dissenting) (quoting Grayned v. City of Rockford, 408 U.S. 104, 108–09 (1972)).

205. *Id.*

principle: the Article I nondelegation doctrine. The idea that the Due Process Clause prohibits the delegation of discretion to the executive, as part of the *Bill of Rights* nondelegation doctrine, is entirely consistent with Justice Thomas's *Dimaya* dissent.

Scholars have expounded upon the parallels. A main observation has been that "[v]ague statutes have the effect of delegating lawmaking authority to the executive."[206] As Michael Mannheimer has written, "the void-for-vagueness doctrine operates as a type of nondelegation doctrine, bolstering the separation of powers by requiring that the lawmaking power be housed in the legislative branch."[207] A recent Note, entitled "Vagueness and Nondelegation," makes the point succinctly: "The void-for-vagueness doctrine and the nondelegation doctrine share an intuitive connection: when Congress drafts vague statutes, it delegates lawmaking authority to courts and the executive."[208] Moreover, two attorneys have urged adoption of the void-for-vagueness standard—which they describe as requiring that criminal laws "(1) be clear enough to provide fair notice and (2) be

206. Nathan S. Chapman & Michael W. McConnell, *Due Process as Separation of Powers*, 121 YALE L.J. 1672, 1806 (2012). Justice Thomas cites this quotation in his *Dimaya* dissent. *See Dimaya*, 138 S. Ct. at 1248 (Thomas, J., dissenting). That said, then-professor (now Judge) Debra Livingston has taken the position that while "broad and overinclusive rules enhance police discretion, . . . a plethora of narrow rules may not meaningfully constrain it, since such rules may or may not be enforced." Debra Livingston, *Police Discretion and the Quality of Life in Public Places: Courts, Communities, and the New Policing*, 97 COLUM. L. REV. 551, 618 (1997).

207. Michael J. Zydney Mannheimer, *Vagueness as Impossibility*, 98 TEX. L. REV. 1049, 1055 (2020); *see also* Guyora Binder & Brenner Fissell, *A Political Interpretation of Vagueness Doctrine*, 2019 U. ILL. L. REV. 1547, 1548–49 (describing the "underlying concern" in the vagueness cases not in explicit nondelegation terms, but as a belief that "discretion allows executive officials to make determinations about what should be punished, but such determinations should only be made by elected legislatures"); Forde-Mazrui, *supra* note 107, at 1500 ("[C]entral to the rule of law is the principle that specificity in legal rules serves to constrain the discretion exercised by those charged with their enforcement. This principle has been constitutionalized by the courts, through the void-for-vagueness doctrine, as a safeguard against legislative delegation of excessive discretion to courts and to executive officials and agencies, especially the police.").

208. Ogale, *supra* note 20, at 783. Ogale contends that "there are two vagueness doctrines"—what he calls "Rights-Based Vagueness" (exemplified by cases like *Papachristou*) and "Structure-Based Vagueness" (exemplified by cases like *Dimaya*). *See id.* at 786–87. In Ogale's view, "To the extent that vagueness and nondelegation converge, it is in the context of Structure-Based Vagueness." *Id.* at 787. But as this Article has shown, even the rights-based vagueness cases indicate concerns about delegation. *See supra* notes 198–200.

enacted by elected legislators to ensure democratic legitimacy"—in nondelegation cases "to police noncriminal delegations as well."[209]

B. Tying It All Together

The Bill of Rights nondelegation doctrine emerges. Taken together, the cases discussed in section II.A stand for a coherent rule: several of the amendments in the Bill of Rights protect a right by allocating power between branches of government, and when such a right is at issue, discretion may not be delegated in a way that upsets that allocation. In the First Amendment cases, the Supreme Court has prohibited legislatures from conferring open-ended discretion on executive officials to deny permits for expressive activity.[210] In *Bruen*—a Second Amendment case—the Court prohibited a state legislature from conferring this same sort of discretion on executive officials to deny concealed-carry permits.[211] Meanwhile, the Fourth Amendment particularity-requirement cases prevent *judges* from delegating discretion—via a warrant—to the executive about what to search and seize.[212] And in the Fifth Amendment (and Fourteenth Amendment) vagueness cases, the Court has required that the *legislature*—not the executive—make the relevant policy choices when crafting penal laws.[213]

In the Bill of Rights nondelegation cases, the Court has focused on the *branch* of government that is supposed to exercise a particular discretionary power. If that branch of government delegates that discretion to another branch of government, the Court has declared the delegation *itself*—whether via statute or warrant—to be unconstitutional. This posture comports with one of the key insights of *Whitman v. American Trucking Associations, Inc.*, a nondelegation case discussed earlier in this Article:[214] the proper way to make out a nondelegation challenge is by challenging the underlying delegation, not the action taken pursuant to that delegation.[215] The Court has articulated different, yet similar, rationales for why the mere existence of a delegation

209. Gaziano & Blevins, *supra* note 20, at 45.

210. *See supra* Section II.A.1.

211. *See supra* Section II.A.2.

212. *See supra* Section II.A.3.

213. *See supra* Section II.A.4.

214. *See supra* text accompanying notes 84–87.

215. Whitman v. Am. Trucking Ass'ns, Inc., 531 U.S. 457, 472 (2001).

impermissibly infringes upon individual liberty. For example, in the First Amendment cases, the Court has found that the "lodg[ing of] broad discretion in a public official to permit speech-related activity" itself abridges speech[216] because the fact that "[d]issemination of ideas depends upon the approval of the distributor by the official . . . is [itself] administrative censorship in an extreme form."[217] Meanwhile, in the Fourth Amendment caselaw, the Court has explained that when people are "secure only in the discretion of police officers[,]" people cannot fully enjoy the security and privacy that the Fourth Amendment guarantees.[218]

To that end, note what these cases are *not* primarily about. The issue in the Bill of Rights nondelegation cases is—at least primarily—the fact of delegation of discretion, not the underlying statute's substantive limitation of the right in question or the executive action that violates the right. In the First Amendment cases, the Court has declared licensing regimes unconstitutional because they delegated discretion to the executive, not because the statutorily prescribed regime itself formally favored one viewpoint over another (a classic example of First Amendment-violative legislation). To be sure, *Bruen* did look to the relationship between New York's proper-cause standard and the right guaranteed by the Second Amendment.[219] But Justice Kavanaugh, concurring in *Bruen*, interpreted the majority opinion to say that "New York's outlier may-issue regime is constitutionally problematic *because* it grants open-ended discretion to licensing officials and authorizes licenses only for those applicants who can show some special need apart from self-defense."[220] As Justice Kavanaugh put it, "[t]hose features of New York's regime—the unchanneled discretion for licensing officials and the special-need requirement—in effect deny the right to carry handguns for self-defense to many 'ordinary, law-abiding citizens.'"[221] Similarly, the void-for-vagueness cases condemn the discretion that vague statutes lodge in the executive, without much inquiry into whether the police power of the state

216. Police Dep't of Chi. v. Mosley, 408 U.S. 92, 97 (1972).

217. Largent v. Texas, 318 U.S. 418, 422 (1943).

218. Johnson v. United States, 333 U.S. 10, 13–14 (1948) (describing the Fourth Amendment as "a nullity" in this circumstance).

219. *See* N.Y. State Rifle & Pistol Ass'n v. Bruen, 142 S. Ct. 2111, 2156 (2022).

220. *Id.* at 2161 (Kavanaugh, J., concurring) (emphasis added).

221. *Id.* (quoting *id.* at 2122 (majority opinion)).

would permit the government to proscribe, for example, loitering on a street corner, if it should do so with the requisite specificity.

Also, the cases do not admit of a distinction between what administrative law calls "rulemaking" and "adjudication." The divide between rulemaking and adjudication "is illustrated by [the Supreme] Court's treatment of two related cases under the Due Process Clause":[222] *Londoner v. City & County of Denver*[223] and *Bi-Metallic Investment Co. v. State Board of Equalization*.[224] Between *Londoner* and *Bi-Metallic*, the Court established "[a] foundational rule of due process in administrative law . . . that due process attaches to administrative adjudication, not rulemaking."[225] Describing legislative rulemaking, the Court in *Bi-Metallic* observed that "[w]here a rule of conduct applies to more than a few people it is impracticable that every one should have a direct voice in its adoption. The Constitution does not require all public acts to be done in town meeting or in an assembly of the whole."[226] The distinction matters in the Article I nondelegation context, too. Enforcing Article I's Vesting Clause against the backdrop of the Constitution's structure, the Article I nondelegation cases have typically concerned the possibility of rules of conduct that apply "to more than a few people"—legislation by the executive. Yet courts in Bill of Rights cases appear to have no issue with exercising judicial review whether the delegated discretion manifests as rulemaking (*e.g.*, the promulgation of criminal standards fleshing out a vague criminal statute) or adjudication (*e.g.*, the denial of an individual license to an applicant for a speech permit).

Judicial review is a weighty undertaking.[227] And perhaps the Article I nondelegation doctrine's history shows that the Court is wary about exercising its power of judicial review to enforce the nondelegation principle. But the Bill of Rights nondelegation cases demonstrate that this wariness is not an absolute bar to the exercise

222. United States v. Fla. E. Coast Ry. Co., 410 U.S. 224, 244 (1973).

223. Londoner v. City & Cnty. of Denver, 210 U.S. 373 (1908).

224. Bi-Metallic Inv. Co. v. State Bd. of Equalization, 239 U.S. 441 (1915).

225. Cass R. Sunstein & Adrian Vermeule, *The Morality of Administrative Law*, 131 HARV. L. REV. 1924, 1964 (2018).

226. *Bi-Metallic*, 239 U.S. at 445.

227. *See supra* Section I.B.2; *see also* WILLIAM H. REHNQUIST, THE SUPREME COURT 144 (1987) ("From the time of John Marshall, the Court has said that the authority to declare an act of Congress unconstitutional is the most awesome responsibility that any court could possess, and the authority to do so must be exercised with extraordinary circumspection.").

of judicial review. Granted, the bulk of these cases are about state statutes (or state warrants).[228] Nevertheless, the Court has employed the Bill of Rights nondelegation doctrine to engage in robust constitutional review on a regular basis.

The Bill of Rights nondelegation doctrine trains its fire on the delegation of discretion when that delegation could lead to the infringement of an enumerated right. The doctrine enforces a rule about which branch of government must exercise a certain, discretionary power—and, equally as important when applying the nondelegation doctrine, which branch of government cannot be delegated that power. The Court has applied the doctrine when merely the potential exercise—by the wrong branch of government—of the discretionary power in question would violate the people's constitutionally guaranteed rights of speech; keeping and bearing arms; security in their persons, houses, papers, and effects; or fair notice of what conduct is prohibited. The fact of the doctrine's existence illuminates an anti-delegation principle across Bill of Rights cases.

III. SOME POTENTIAL APPLICATIONS

The identification of a Bill of Rights nondelegation doctrine could have significant consequences for Bill of Rights jurisprudence overall. For the purpose of illustrating the point, this Part touches upon three discrete scenarios in which the Bill of Rights nondelegation doctrine could have an impact. To be sure, this Part merely provides some examples; it does not intend to be an exhaustive summary of the doctrine's potential applications, and this Article does not claim that the doctrine necessarily applies across the entirety of the Bill of Rights. As to the three applications discussed in this Part: First, the doctrine could provide a path forward for judicial scrutiny of certain misleadingly labeled "shall-issue" concealed carry permit jurisdictions in the wake of *Bruen*, even if they operate like shall-issue regimes. Second, the doctrine could supply a framework for understanding whether and how courts should defer to congressional judgments of what is "reasonable" for the purposes of the Fourth Amendment. And third, the doctrine could solve what this Article calls the "*Jarkesy*

228. *But see, e.g.*, Sessions v. Dimaya, 138 S. Ct. 1204 (2018); Johnson v. United States, 576 U.S. 591 (2015).

problem," looking to a recent Fifth Circuit decision now before the Supreme Court—*Jarkesy v. SEC*[229]—that attempted to square the Article I nondelegation doctrine with an issue of discretionary power to violate individuals' jury trial rights.

A. Perhaps Misleadingly Labeled "Shall-Issue" Concealed Carry Permit Jurisdictions

In *Bruen*, the Court confronted a discretionary permitting regime for concealed-carry permits that clearly violated the Bill of Rights nondelegation doctrine. With no standards by which the licensing authority was directed to determine "proper cause," the permitting scheme delegated "open-ended discretion to licensing officials."[230] The Court separated the different concealed-carry permitting regimes of the U.S. states into three buckets. First, the Court found that 43 states were "shall issue" jurisdictions, meaning that "authorities must issue concealed-carry licenses whenever applicants satisfy certain threshold requirements, without granting licensing officials discretion to deny licenses based on a perceived lack of need or suitability."[231] Second, the Court explained that six states (including New York) and the District of Columbia operated "may issue" regimes, "under which authorities have discretion to deny concealed-carry licenses even when the applicant satisfies the statutory criteria, usually because the applicant has not demonstrated cause or suitability for the relevant license."[232] Third and finally, the Court noted that one state—Vermont—had "no permitting system for the concealed carry of handguns."[233] A reasonable inference to draw from *Bruen* is that the second category of jurisdictions is constitutionally dubious while the first and third categories are likely fine.

But not all of the 43 purportedly "shall-issue" regimes are the same, and even some of those states' permitting schemes might violate the Bill of Rights nondelegation doctrine. In the first footnote of *Bruen*, the Court explained that "[t]hree States—Connecticut, Delaware, and Rhode Island—have discretionary

229. Jarkesy v. SEC, 34 F.4th 446 (5th Cir. 2022).

230. N.Y. State Rifle & Pistol Ass'n v. Bruen, 142 S. Ct. 2111, 2161 (2022) (Kavanaugh, J., concurring).

231. *Id.* at 2123 (majority opinion).

232. *Id.* at 2123–24.

233. *Id.* at 2123 n.1.

criteria but appear to operate like 'shall issue' jurisdictions."[234] Recall the key insight of *Whitman*: in an Article I nondelegation challenge, the court evaluates the underlying statute.[235] Under *Whitman*, the very fact of delegation cannot be cured by an executive's narrowing interpretation of the discretion-delegating statute. Thus, the question for a court applying the Bill of Rights nondelegation doctrine is not whether the regime "appear[s] to operate like [a] 'shall issue' jurisdiction," but whether the underlying statute itself delegates unfettered discretion to the executive in a way that the doctrine prohibits.

A few basic legal principles are helpful in framing this inquiry. To start, federal courts "are bound by the construction" that state courts give to their own states' statutes.[236] Moreover, federal courts may accept a state supreme court's "narrowing of a state statute" "to avoid constitutional infirmities."[237] For these reasons, a state supreme court's discretion-cabining construction of a discretion-granting concealed-carry permitting regime likely cannot be disturbed by a federal court.[238]

Turning to the Connecticut, Delaware, and Rhode Island laws, the Bill of Rights nondelegation doctrine may change the way that federal courts should think about at least one of these states' concealed-carry permitting schemes. To start, Connecticut and Rhode Island are likely properly classified as shall-issue jurisdictions. The *Bruen* Court noted how both the Connecticut and Rhode Island courts have interpreted their concealed-carry permitting schemes to narrow discretion in such a way that does not present a constitutional problem.[239] Whether those interpretations are correct is a separate question, but that question

234. *Id.*

235. *See supra* Section I.B.2.iv.

236. R.A.V. v. St. Paul, 505 U.S. 377, 381 (1992).

237. Iancu v. Brunetti, 139 S. Ct. 2294, 2312 (2019) (Sotomayor, J., concurring in part and dissenting in part) (citing Chaplinsky v. New Hampshire, 315 U.S. 568 (1942)); *see also* Bell v. Cone, 543 U.S. 447, 456–60 (2005) (accepting a state court's narrowing construction of a state statute's "aggravating circumstance" to cure a vagueness problem).

238. This Article assumes that the Bill of Rights nondelegation doctrine—even against the backdrop of the Fourteenth Amendment's incorporation of the Bill of Rights against the states—does not override the ordinary rule that federal courts must accept state supreme courts' constructions of state law when those constructions cure delegation issues.

239. *See Bruen*, 142 S. Ct. at 2123 n.1.

is not one that the Bill of Rights nondelegation doctrine would have anything to say about—at least not in the federal courts.

Yet Delaware is different in kind. Rather than pointing to a Delaware court's narrowing construction of the permitting regime, the Court noted that as of its decision in *Bruen*, "the State ha[d] thus far processed 5,680 license applications and renewals in fiscal year 2022 and ha[d] denied only 112."[240] Relying on this justification, however, presents a *Whitman* problem. The fact that the government has prudently exercised improperly delegated discretion does not obviate what would otherwise be a Bill of Rights nondelegation issue.

Dissenting in *Bruen*, Justice Breyer noted an inconsistency in the Court's classification of the different regimes. Justice Breyer questioned why the Court deemed Connecticut, Delaware, and Rhode Island to be shall-issue jurisdictions when it recognized them as having may-issue statutory criteria.[241] As Justice Breyer explained, "[T]hese three States demonstrate [that] the line between 'may issue' and 'shall issue' regimes is not as clear cut as the Court suggests, and that line depends at least in part on how statutory discretion is applied in practice."[242] Particularly as to Delaware, Justice Breyer is correct—and the Bill of Rights nondelegation doctrine provides the proper framework for understanding why. Whether Delaware operates in *practice* like a shall-issue jurisdiction is immaterial. The very fact of delegated discretion likely renders it a may-issue jurisdiction. For this reason, even in light of the way Delaware administers its concealed-carry licensing regime, *Bruen* seems to indicate that the scheme is unconstitutional when applying the Bill of Rights nondelegation doctrine.

B. Congressional Determinations of Fourth Amendment Reasonableness

The particularity requirement for warrants is not the only aspect of the Fourth Amendment on which the Bill of Rights nondelegation doctrine might bear. And the executive is not the only branch of government to which the doctrine would prevent delegation. The Fourth Amendment provides that "[t]he right of the people to be secure in their persons, houses, papers, and effects,

240. *Id.*

241. *See id.* at 2172 (Breyer, J., dissenting).

242. *Id.*

against unreasonable searches and seizures, shall not be violated[.]"[243] The question of reasonableness usually turns on the question of whether the government has obtained a warrant.[244] But courts sometimes determine what is "reasonable" for Fourth Amendment purposes with reference to the judgment of a legislature.[245]

In *United States v. Watson*, the Supreme Court permitted introduction of evidence obtained pursuant to an arrest carried out by a federal postal inspector, despite the fact that the government had not obtained a warrant for the arrest.[246] A federal statute had authorized such arrests.[247] In the Court's view, that statute "represent[ed] a judgment by Congress that it is not unreasonable under the Fourth Amendment for postal inspectors to arrest without a warrant provided they have probable cause to do so."[248] Quoting *United States v. Di Re*, the Court observed that it "should be reluctant to decide that a search . . . authorized by Congress was unreasonable and that the Act was therefore unconstitutional."[249] The Court noted that securing a warrant in advance of an arrest was ordinarily preferable.[250] But it "decline[d] to transform this judicial preference into a constitutional rule when the judgment of the Nation and Congress has for so long been to authorize warrantless public arrests on probable cause[.]"[251]

The relationship between wiretapping and the Fourth Amendment provides another example of judicial deference to congressional judgments of reasonableness. Initially, the Court in *Olmstead v. United States* held that wiretapping did not violate the Fourth Amendment.[252] Nearly forty years later, the Court in *Katz v.*

243. U.S. CONST. amend. IV.

244. *See* City of Ontario v. Quon, 560 U.S. 746, 760 (2010) ("Although as a general matter, warrantless searches are per se unreasonable under the Fourth Amendment, there are a few specifically established and well-delineated exceptions to that general rule." (internal quotation marks omitted)).

245. *See, e.g.*, United States v. Watson, 423 U.S. 411, 416–17 (1976); United States v. Di Re, 332 U.S. 581, 585 (1948).

246. *See Watson*, 423 U.S. at 423–24.

247. *See id.* at 414–15.

248. *Id.* at 415.

249. *Id.* at 416 (quoting United States v. Di Re, 332 U.S. 581, 585 (1948)).

250. *See id.* at 423.

251. *Id.*

252. *See* Olmstead v. United States, 277 U.S. 438 (1928).

United States took a different tack, determining that "[t]he Government's activities in electronically listening to and recording the petitioner's words violated the privacy upon which he justifiably relied while using the telephone booth and thus constituted a 'search and seizure' within the meaning of the Fourth Amendment."[253] Then Congress stepped in, enacting Title III of the Omnibus Crime Control and Safe Streets Act of 1968.[254] As Justice Alito has pointed out: "Since that time, electronic surveillance has been governed primarily, not by decisions of [the] Court, but by the statute, which authorizes, but imposes detailed restrictions on, electronic surveillance."[255] Some Justices take the position that when it comes to the Fourth Amendment, "[l]egislatures, elected by the people, are in a better position than [judges] are to assess and respond to the changes that have already occurred and those that almost certainly will take place in the future."[256] That view, if taken to its logical conclusion—judicial deference to legislative judgments about reasonableness—would present a constitutional difficulty when considered through the lens of the Bill of Rights nondelegation doctrine.

Courts are the proper determiners—in the first instance—of whether a search violates the Fourth Amendment. Legislatures are certainly more democratically accountable than courts are. And no one doubts that "legislatures (or agencies) can . . . create additional protections" above that which the courts have determined the

253. Katz v. United States, 389 U.S. 347, 353 (1967).

254. *See* Riley v. California, 573 U.S. 373, 408 (2014) (Alito, J., concurring in part and concurring in the judgment).

255. *Id.*; *see also* United States v. Jones, 565 U.S. 400, 427–28 (Alito, J., concurring in the judgment) ("After *Katz*, Congress did not leave it to the courts to develop a body of Fourth Amendment case law governing that complex subject. Instead, Congress promptly enacted a comprehensive statute, and since that time, the regulation of wiretapping has been governed primarily by statute and not by case law." (first citing 18 U.S.C. §§ 2510–22 (2006); and then citing Orin S. Kerr, *The Fourth Amendment and New Technologies: Constitutional Myths and the Case for Caution*, 102 MICH. L. REV. 801, 850–51 (2004)).

256. *Riley*, 573 U.S. at 408 (Alito, J., concurring in the judgment); *see also* Carpenter v. United States, 138 S. Ct. 2206, 2233 (Kennedy, J., dissenting) ("In § 2703(d) Congress weighed the privacy interests at stake and imposed a judicial check to prevent executive overreach. The Court should be wary of upsetting that legislative balance and erecting constitutional barriers that foreclose further legislative instructions. . . . The Court's decision runs roughshod over the mechanism Congress put in place to govern the acquisition of cell-site records and closes off further legislative debate on these issues." (citing City of Ontario v. Quon, 130 S. Ct. 2619 (2010))).

Fourth Amendment's floor to be.[257] Indeed, Title III may in fact represent additional protections. Moreover, reference to a legislature's judgment can provide evidence of what the society finds to be reasonable. But the issue comes when a court defers to a statute's reasonableness determination if that statute goes below the floor of protection that the court would otherwise believe the Fourth Amendment secures.[258] In effect, wholesale judicial deference to a legislature's reasonableness determination works a reverse delegation of discretion—*from* the courts *to* the legislature—and contravenes the cardinal constitutional rule that "fundamental rights may not be submitted to vote" because "they depend on the outcome of no elections."[259] Such deference to the legislature violates the Bill of Rights nondelegation doctrine.

In practice, deference to congressional judgments about reasonableness prevents the courts from undertaking an independent inquiry into the Fourth Amendment's floor. That independent inquiry guards against a legislature's recalibration of the balance that the Framers already struck with respect to the Fourth Amendment's protections. The inquiry entails the exercise of discretion—determining what is "reasonable" implicates a variety of considerations with no constraining principle. As discussed earlier in this Article, that discretion is dangerous. And in our system, the court must be the one exercising the discretion in this particular context.

C. Solving the Jarkesy *Problem*

The Fifth Circuit recently detonated an administrative and constitutional law bomb in *Jarkesy v. SEC*.[260] There, a Fifth Circuit panel picked apart various aspects of a Securities and Exchange Commission (SEC) adjudicatory scheme on constitutional grounds,

257. Matthew Tokson, *The Aftermath of* Carpenter*: An Empirical Study of Fourth Amendment Law, 2018–2021*, 135 HARV. L. REV. 1790, 1851 (2022).

258. *Cf.* Silver, *supra* note 2121, at 1257 (describing how state "[a]ppellate courts are uniformly skeptical when a trial court farms out its decisionmaking powers to experts[,]" including reference to a Maine case in which the court "invalidated a parental rights order that 'contact between the father and the older child shall resume "as therapeutically recommended"' . . . because, while 'the court can consider a therapist's opinion' in determining parental rights, 'the court cannot make the visitation outcome dependent upon that opinion'" (quoting *In re* Children of Richard E., 227 A.3d 159, 169 (Me. 2020)).

259. W. Va. State Bd. of Educ. v. Barnette, 319 U.S. 624, 638 (1943).

260. Jarkesy v. SEC, 34 F.4th 446 (5th Cir. 2022).

including invocation of the Article I nondelegation doctrine.[261] The Supreme Court has granted certiorari in the case.[262]

In the early 2010s, after an investigation, the SEC determined that George Jarkesy had probably committed securities fraud.[263] The agency then decided it would bring charges against Jarkesy. The steps here are important: the SEC (1) investigated Jarkesy, (2) concluded that he likely violated multiple federal securities laws, and (3) decided to bring charges. But before the SEC could bring those charges at step three, it still had one more thing to do: decide the *forum* in which it wanted to bring an enforcement action against Jarkesy. Under the Dodd-Frank Act, the SEC could choose to bring an enforcement action either within the agency ("in-house")—with no jury afforded to the subject of the enforcement action—or in an Article III federal court.[264]

The SEC has publicly stated that "when the misconduct warrants it, the Commission will bring both proceedings."[265] In-house adjudication occurs before an administrative law judge (ALJ). This adjudicatory regime is often far more efficient and, the numbers show, slanted in the agency's favor.[266] The ALJs are themselves SEC employees.[267] Elizabeth Wang explains the difference between federal court adjudication and in-house adjudication well:

> In federal court, defendants have access to a jury trial, independent judges, and deposition "testimony [that] is subjected to the Federal Rules of Evidence." Alternatively, administrative proceedings are conducted before an ALJ, where there is no jury,

261. *See id.* at 449–50.

262. SEC v. Jarkesy, 143 S. Ct. 2688 (2023) (mem.).

263. *Jarkesy*, 34 F.4th at 450.

264. *See id.* at 455 (citing 15 U.S.C. § 78u-2(a)).

265. *How Investigations Work*, U.S. SECURITIES & EXCHANGE COMM'N, https://www.sec.gov/enforce/how-investigations-work.html (last visited Oct. 4, 2023).

266. *See* Jean Eaglesham, *Fairness of SEC Judges Is in Spotlight*, WALL ST. J. (Nov. 22, 2015, 9:25 PM), https://www.wsj.com/articles/fairness-of-sec-judges-is-in-spotlight-1448236970.

267. *See id. But cf. id.* (quoting an SEC ALJ as saying, "The SEC can't fire us, decide our pay or grade our performance . . . . There's nothing the SEC can do to influence us and they don't try to").

> discovery is restricted, hearings proceed on a rapid schedule, and the Federal Rules of Evidence do not apply.[268]

For the SEC, bringing the charges before an ALJ saves time, yields a high rate of success, and gets the case before an expert adjudicator whose primary role is to hear cases about securities law violations (as opposed to generalist Article III judges). It is no wonder, then, that "[t]he SEC has recently leaned more heavily on its in-house tribunal."[269] Naturally, the SEC brought its charges against Jarkesy in-house.[270]

In *Jarkesy*, the Fifth Circuit held—among other things—that the forum-selection provision of the Dodd-Frank Act violated the Article I nondelegation doctrine.[271] In the statute, Congress provided no guidance for how the SEC was to choose between these two options. In the Fifth Circuit's telling, "Congress gave the SEC a significant legislative power by failing to provide it with an intelligible principle to guide its use of the delegated power."[272] That legislative power, according to the Fifth Circuit, was "the unfettered authority to choose whether to bring enforcement actions in Article III courts or within the agency."[273]

Jonathan Adler disagreed with the court's Article I nondelegation holding. He put it the following way: "The delegated power at issue is the SEC's authority to make case-by-case decisions about how to enforce the securities laws against individual regulated entities. This is not legislative power."[274] Rather, Adler wrote, "[t]his is the sort of prosecutorial discretion that lies at the core of executive authority. And because this is not legislative power, no 'intelligible principle' is required."[275] The Fifth Circuit had written that

268. Elizabeth Wang, Comment, Lucia v. SEC*: The Debate and Decision Concerning the Constitutionality of SEC Administrative Proceedings*, 50 LOY. L.A. L. REV. 867, 870 (2017) (quoting Joseph Grundfest, *Fair or Foul? SEC Administrative Proceedings and Prospects for Reform Through Removal Legislation* 1 (Rock Ctr. for Corp. Governance at Stanford Univ., Working Paper No. 212, 2015; then citing *id.* at 3–6)).

269. Eaglesham, *supra* note 266.

270. Jarkesy v. SEC, 34 F.4th 446, 449 (5th Cir. 2022).

271. *See id.* at 451.

272. *Id.* at 459.

273. *Id.*

274. Jonathan H. Adler, *The Good, the Bad, and the Ugly of* Jarkesy v. SEC, REASON: VOLOKH CONSPIRACY (Aug. 17, 2022, 6:10 PM), https://reason.com/volokh/2022/08/17/the-good-the-bad-and-the-ugly-of-jarkesy-v-sec.

275. *Id.*

"[g]overnment actions are 'legislative' if they have 'the purpose and effect of altering the legal rights, duties and relations of persons . . . outside the legislative branch.'"[276] Adler countered that this definition "doesn't do the work the Fifth Circuit wants it to. Jarkesy's rights in an Article III court and in an administrative proceeding are what they are under the Constitution and relevant statutes. The SEC did not alter these rights. It merely chose how to enforce the laws Congress enacted."[277]

Adler is half-right. The delegation to the SEC to determine the forum in which to bring individual enforcement actions is significantly different than what the Article I nondelegation doctrine has traditionally condemned as delegation of *legislative* power. But the forum determination is not the same thing as prosecutorial discretion. When a prosecutor chooses whether to litigate in an Article III court or before an agency, that choice is different than the choice of the statute under which to prosecute or the choice of whether to prosecute at all. Ordinarily, as described above, an agency decides two things when bringing an enforcement action: (1) what statute—or implementing regulation—the alleged offender violated, and (2) whether to bring the action. But the Dodd-Frank Act added a third step to this decision-making process for the SEC: the question of the forum in which to bring the action (in a federal court or before the agency itself). And as the cases demonstrate, the Bill of Rights nondelegation doctrine does not admit of a distinction between legislative rulemaking and adjudication when adjudication requires discretion and can lead to a rights violation.[278] So even though the agency would exercise discretion on an individual basis with respect to individual subjects of enforcement, the underlying statute has a nondelegation problem in the same way that the speech licensing regimes had nondelegation problems in the cases discussed earlier in this Article.

Choosing in-house adjudication has real consequences. One such consequence has a Bill of Rights nexus: in-house adjudication provides no jury, potentially contravening the Seventh Amendment's guarantee that "[i]n suits at common law, where the value in

276. *Jarkesy*, 34 F.4th at 461 (alteration in original) (quoting INS v. Chadha, 462 U.S. 919, 952 (1983)).

277. Adler, *supra* note 274.

278. *See supra* notes 222–226.

controversy shall exceed twenty dollars, the right of trial by jury shall be preserved."[279] Indeed, the Fifth Circuit in *Jarkesy* also held that the SEC's in-house adjudicatory scheme violated Jarkesy's Seventh Amendment right to a civil jury.[280]

The better way to think about the delegation problem in *Jarkesy* is with reference to the Bill of Rights nondelegation doctrine. The problem is an agency's ability to decide—in its unfettered discretion—to litigate in a forum in which a subject of enforcement gets no jury. Under the Bill of Rights nondelegation doctrine, Congress cannot delegate this sort of discretion to an administrative agency when an enumerated right—here, the right to a trial by jury—is at stake. Moreover, applying *Whitman*, the SEC could not itself cure the delegation problem by setting forth limits on its own discretion.[281] Assuming the in-house adjudication would violate the Seventh Amendment, the underlying statute thus likely violates not the Article I nondelegation doctrine but rather the Bill of Rights nondelegation doctrine.

Jarkesy is therefore a Bill of Rights nondelegation case. The problem is not just the statute as applied to Jarkesy; under the Court's precedents, the whole forum-selection regime is vulnerable to a facial challenge because it impermissibly lodges discretion in the SEC with respect to an enumerated right. Even if the SEC never exercised its discretion to bring an enforcement action in the administrative forum (not the state of the world, to be sure), the law would still contravene a principle of nondelegation that has permeated the Court's jurisprudence. And even though the SEC is probably not exercising a *legislative* power when determining—on an individual basis—whether to bring particular enforcement actions in-house or in federal court, the nondelegation doctrine (a

279. U.S. CONST. amend. VII. Determining that the no-jury scheme of the administrative tribunal violates the Bill of Rights nondelegation doctrine likely requires an antecedent determination that having to submit to administrative adjudication would violate one's Seventh Amendment right to a trial by jury. For an argument that the Constitution prohibits the juryless tribunals that have become a hallmark of administrative adjudication, see Richard Lorren Jolly, *The Administrative State's Jury Problem*, 98 WASH. L. REV. (forthcoming 2023).

280. *Jarkesy*, 34 F.4th at 465.

281. The SEC had done exactly this, having "issued internal guidance on the selection between administrative and civil proceedings." Kenneth Oshita, *Home Court Advantage? The SEC and Administrative Fairness*, 90 S. CAL. L. REV. 879, 887 (2017); *see also* David Zaring, *Enforcement Discretion at the SEC*, 94 TEX. L. REV. 1155, 1207 (2016).

specific kind of nondelegation doctrine) still has something to say about the statute conferring that discretion.

IV. CAVEATS

This Part addresses some counterarguments and clarifies this Article's thesis. Three points are important. First, the Bill of Rights nondelegation doctrine and the Article I nondelegation doctrine differ in important ways. Second, the discretion at which the Bill of Rights nondelegation doctrine takes aim is different from prosecutorial discretion, which the doctrine does not address. Third, the Bill of Rights nondelegation doctrine is far from the only way in which courts enforce the Bill of Rights' protections, and it is not necessarily enforced across the entirety of the Bill of Rights. Nevertheless, reading the Bill of Rights nondelegation cases in relation to one another illuminates how the Supreme Court has given teeth to the Bill of Rights.

A. The Article I Nondelegation Doctrine vs. the Bill of Rights Nondelegation Doctrine

The Bill of Rights nondelegation doctrine is not the Article I nondelegation doctrine. Important, substantive differences exist between the two doctrines, which both fall under the umbrella of the "nondelegation doctrine." The Article I nondelegation doctrine prohibits Congress from delegating—to the executive—any of the *legislative powers* with which the Constitution has vested Congress. Meanwhile, the Bill of Rights nondelegation doctrine prohibits the delegation of *discretion* to a branch of government when it would upset the Constitution's allocation of power concerning an enumerated right. In the end, the most meaningful parallel is the bar on delegation itself, but the two doctrines are not the same.[282]

282. Understanding the Article I nondelegation doctrine as a distinct aspect of the nondelegation doctrine may shed some light on the relevance of certain legal materials to the question of whether the Article I nondelegation doctrine is consistent with the original meaning of the Constitution. *See, e.g.*, Eli Nachmany, *The Irrelevance of the Northwest Ordinance Example to the Debate About Originalism and the Nondelegation Doctrine*, 2022 U. ILL. L. REV. ONLINE 17 (2022) (arguing that the broad delegation of lawmaking authority in the Northwest Ordinance sheds no light on the original meaning of the Article I nondelegation doctrine because Congress enacted the ordinance pursuant to its power under Article IV as opposed to an Article I power).

Legislative power and discretion are similar. The term "legislative power" connotes the discretionary power to prescribe—subject only to the constraints imposed by the Constitution[283]—the rules by which conduct is ordered in a given society. "Discretion" itself has a more particular definition. The relevant definition in Black's Law Dictionary is "[f]reedom in the exercise of judgment; the power of free decision-making."[284] But this freedom, in the hands of the executive, invites the arbitrary exercise of will. The Bill of Rights nondelegation doctrine merely takes this observation, applies it when an enumerated right is at stake (whether as a result of legislative rulemaking or administrative adjudication), and safeguards this right by prohibiting the delegation of discretion to the wrong branch of government.

For that reason, the key takeaway is that the Bill of Rights nondelegation doctrine is just as much about delegation as is the Article I nondelegation doctrine.[285] In the First, Second, and Fifth Amendment cases, the delegation is clear: a legislature has delegated discretion to an executive official. That discretion could manifest as the power to deny a permit for expressive activity, to deny a permit to carry a concealed firearm, or to enforce a vague criminal ordinance. The Fourth Amendment cases are a bit trickier to analogize, but they too are about delegation. Here, the delegation is from the judicial officer to the police. Under the Constitution, the judicial magistrate is supposed to be the one who exercises the discretionary power to determine whether the warrant describes with particularity "the place to be searched, and the persons or things to be seized."[286] Just as the Article I nondelegation doctrine

283. For a particularly strong version of this argument, see generally Randy E. Barnett & Evan D. Bernick, *No Arbitrary Power: An Originalist Theory of the Due Process of Law*, 60 WM. & MARY L. REV. 1599 (2019) (advancing an interpretation of the Due Process Clause that requires good-faith exercise of legislators' discretionary powers in a manner actually calculated to achieve constitutionally proper ends).

284. *Discretion*, BLACK'S LAW DICTIONARY (11th ed. 2019); *see also Legal Theory Lexicon 091: Discretion*, LEGAL THEORY LEXICON (Oct. 22, 2022), https://lsolum.typepad.com/legal_theory_lexicon/2019/03/legal-theory-lexicon-091-discretion.html.

285. *Cf.* Indus. Union Dep't, AFL-CIO v. Am. Petroleum Inst. (*Benzene Case*), 448 U.S. 607, 685–86 (1980) (Rehnquist, J., concurring) (describing the Article I nondelegation doctrine as "ensur[ing] that courts charged with reviewing the exercise of *delegated legislative discretion* will be able to test that exercise against ascertainable standards" (emphasis added)).

286. U.S. CONST. amend. IV. To be sure, this discretionary power is *itself* subject to a standard: the warrant may only issue "upon probable cause, supported by Oath or affirmation." *Id.*

prohibits Congress from delegating certain legislative powers to the executive, the Bill of Rights nondelegation doctrine prohibits the delegation of unfettered discretion when an enumerated right is at stake.

Still, separating the Article I nondelegation doctrine from the Bill of Rights nondelegation doctrine illuminates a deeper truth: the nondelegation doctrine is about more than Article I of the Constitution. The nondelegation doctrine is an umbrella term for at least two doctrines of constitutional law that can coexist. Whether the Supreme Court has applied the Article I nondelegation doctrine in the years since 1935 bears only on the continued vitality of that version of the nondelegation doctrine. As the cases demonstrate, the Bill of Rights nondelegation doctrine is alive and well.

B. Prosecutorial Discretion vs. Delegated Discretion

The Bill of Rights nondelegation doctrine does not disallow *all* executive discretion. The doctrine leaves prosecutorial discretion—the classic example of permissible executive discretion—undisturbed. To see the point here, one must understand the difference between prosecutorial discretion and the other sort of discretion at which the doctrine takes aim.

Prosecutorial discretion is "the power of the Executive to determine how, when, and whether to initiate and pursue enforcement proceedings."[287] Given the executive's need to allocate limited prosecutorial resources effectively, the traditional justification for the prosecutor's exercise of discretion in enforcing the law is that the discretion is a necessary corollary to the discharge

287. Peter L. Markowitz, *Prosecutorial Discretion Power at Its Zenith: The Power to Protect Liberty*, 97 B.U. L. REV. 489, 490 (2017); *see also* Andrew Kent, Ethan J. Leib & Jed Handelsman Shugerman, *Faithful Execution and Article II*, 132 HARV. L. REV. 2111, 2115 (2019) (discussing the traditional understanding of federal prosecutorial discretion's constitutional textual source); Josh Bowers, *Legal Guilt, Normative Innocence, and the Equitable Decision Not to Prosecute*, 110 COLUM. L. REV. 1655, 1659 (2010) ("[W]hen it comes to critical determinations of normative blameworthiness in petty public order cases, prosecutors enjoy almost unbridled equitable discretion."); Memorandum for the Commissioner from Bo Cooper, INS General Counsel, on INS Exercise of Prosecutorial Discretion 2 (July 11, 2000), https://niwaplibrary.wcl.american.edu/wp-content/uploads/2015/IMM-Gov-ProsDisc-07.11.00.pdf ("Although prosecutorial discretion is sometimes viewed solely as the decision of a prosecutor whether or not to bring charges against an individual, the term also can apply to a broad spectrum of discretionary enforcement decisions taken by a law enforcement agency . . . .").

of the executive's duty.[288] The Bill of Rights nondelegation doctrine has nothing to say about prosecutorial discretion, largely because the executive does not exercise such discretion pursuant to a delegation. Rather, the executive possesses an "inherent" prosecutorial discretion, yielding only to a "clear and specific" statutory limitation and ordinarily not subject to judicial review.[289] Thus, the legislature may proscribe certain conduct (at times in violation of an enumerated right in the Bill of Rights), but the executive's discretion about how to allocate resources in enforcing that proscription is not a problem of delegation.

The discretion at issue in the Bill of Rights nondelegation cases is of a different kind. In these cases, the discretion goes to the nature of the law itself. Suppose that a prosecutor's office has a readily prosecutable case against a suspect thought to have committed murder. The law of murder is clear, and the prosecutor can potentially make the case that this suspect committed murder. Nevertheless, given how difficult it might be to collect the evidence, the prosecutor declines to prosecute. That is prosecutorial discretion. By contrast, suppose that the legislature had delegated to the prosecutor the power to define *what* the law of murder was, such that the prosecutor could decide whether the law should encompass the conduct in which the suspect had engaged. Here, the prosecutor has not been endowed with *prosecutorial* discretion. Rather, Congress has endowed him with a *lawmaking* discretion. The Bill of Rights nondelegation doctrine polices the latter kind of power grants—including conferrals of the power to determine whether a given individual has satisfied the necessary criteria to obtain a permit to speak or to carry a gun (an administrative adjudication). Understanding this distinction helps to clarify *Jarkesy*

288. *See generally* Stephanos Bibas, *The Need for Prosecutorial Discretion*, 19 TEMP. POL. & CIVIL RTS. L. REV. 369 (2010).

289. Cooper, *supra* note 287, at 8. To be sure, prosecutorial discretion is still subject to constitutional constraints. *See, e.g.*, Yick Wo v. Hopkins, 118 U.S. 356, 373–74 (1886) ("Though the law itself be fair on its face, and impartial in appearance, yet, if it is applied and administered by public authority with an evil eye and an unequal hand, so as practically to make unjust and illegal discriminations between persons in similar circumstances, material to their rights, the denial of equal justice is still within the prohibition of the Constitution."); *cf.* United States v. Texas, 577 U.S. 1101 (2016) (mem.) (granting a petition for a writ of certiorari and directing the parties to brief the question whether an executive policy of non-enforcement of immigration law as to a particular subset of aliens "violates the Take Care Clause of the Constitution" before affirming the judgment of the lower court by an equally divided Court at 579 U.S. 547, 548 (2016)).

as a Bill of Rights nondelegation case masquerading as an ordinary Article I nondelegation case.

C. Bill of Rights Jurisprudence vs. Bill of Rights Nondelegation

Courts enforce the Bill of Rights in a variety of ways. The Bill of Rights nondelegation doctrine is merely one such way. The vast majority of Bill of Rights cases focus not on the delegation of discretion to the executive but on the rights infringements themselves. Thus, the Bill of Rights nondelegation doctrine is best understood as existing within a broader framework of protections that the Bill of Rights guarantees the people. Moreover, the Bill of Rights nondelegation doctrine does not necessarily apply across the board.

The Supreme Court is solicitous of individual rights claims when those rights are enumerated in the Bill of Rights. To take one example, the Court has routinely declared state and federal laws to be unconstitutionally violative of the First Amendment. From the State of Texas's anti-flag burning statute in *Texas v. Johnson*[290] to the federal Stolen Valor Act in *United States v. Alvarez*,[291] the Court has exercised its power of judicial review many times when a law prohibits that which the Court believes the First Amendment protects. The same is true when the Court hears challenges under the Second Amendment.[292]

In these cases, the Court has confronted claims that either a "ban" or a limitation on this or that conduct impermissibly infringes on an enumerated right. When litigants make this point, the Court listens. While the Court often engages in "some type of means-end scrutiny" when evaluating these claims,[293] the cases demonstrate that the Court has frequently found in favor of challengers. These claims are different than those that undergird the Bill of Rights nondelegation doctrine. For example, in both *McDonald v. City of Chicago* and *District of Columbia v. Heller*, the challenges under the Second Amendment were to laws that affirmatively banned handguns.[294] Affirmative bans and limitations

290. Texas v. Johnson, 491 U.S. 397 (1989).

291. United States v. Alvarez, 567 U.S. 709 (2012).

292. *See, e.g.*, McDonald v. City of Chicago, 561 U.S. 742 (2010); District of Columbia v. Heller, 554 U.S. 570 (2008).

293. Joseph Blocher, *Bans*, 129 YALE L.J. 308, 310 (2019).

294. *See McDonald*, 561 U.S. 742; *Heller*, 554 U.S. 570.

are different—and perhaps facially more severe—than are mere grants of discretion to deny certain rights to certain individuals. Nevertheless, the Bill of Rights nondelegation doctrine is a component part of the Court's Bill of Rights jurisprudence. Still, this Article does not purport to claim that the doctrine has influenced the jurisprudence of every nook and cranny of the Bill of Rights—at least not yet.

CONCLUSION

The nondelegation doctrine is about more than congressional delegation of legislative power to the executive branch. While those delegations may pose problems under the Article I nondelegation doctrine, the Supreme Court has long been applying another type of nondelegation doctrine in Bill of Rights cases. In cases involving permitting regimes for speech and guns, the Bill of Rights nondelegation doctrine has disfavored the delegation of license-granting discretion to the executive. These cases typically declare that the entire permitting regime is unconstitutional because it confers too much discretion on the wrong governmental actor. When it comes to the First and Second Amendment, discretion about whether and when to grant permits must reside in the legislature—not the executive. When that discretion is delegated, the Bill of Rights nondelegation doctrine has come into play. The Fourth and Fifth Amendment caselaw also sounds nondelegation notes. The Fourth Amendment's particularity requirement for warrants is a rule of nondelegation; it prohibits the delegation of discretion from the courts to the police about what to search or seize. And the Fifth Amendment's void-for-vagueness doctrine prevents the delegation of penal lawmaking power from the legislature to the executive.

The Bill of Rights nondelegation doctrine has several other potential applications—this Article discusses three. First, taking the logic of *Bruen*, courts might determine that other concealed-carry permitting regimes are unconstitutional, even if the executive has purported to limit its own discretion in carrying out the scheme. Second, courts should be careful about deference to the legislature's determination of "reasonableness" for the purposes of the Fourth Amendment's protections. Third, seeing *Jarkesy v. SEC* as a Bill of Rights nondelegation case might clarify the Fifth Circuit's nondelegation holding in its panel opinion. These potential

applications are not an exhaustive list, but they demonstrate the ways that recognizing the doctrine could change our law.

Scholars and jurists have analyzed, applied, called for the revival of, and written obituaries for something called the "nondelegation doctrine." Often, they are talking about the Article I nondelegation doctrine—a component of a broader nondelegation doctrine. As this Article demonstrates, the nondelegation doctrine also has a Bill of Rights component. Recognizing the Bill of Rights nondelegation doctrine could therefore help focus the nondelegation debate, clarifying the scope of originalist inquiry into the nondelegation doctrine's historical pedigree.

The nondelegation doctrine is not dead. Indeed, it has been alive—at least a form of it has been alive—at the Supreme Court for many years. The Bill of Rights cases bear this out. From the First and Second Amendments to the Fourth and Fifth Amendments, the Court has frequently prohibited the delegation of discretion to violate enumerated rights.

Balance in the Basin

Casey Lee McClellan[*]

The National Environmental Policy Act (NEPA) changed the way land managers and users interact with public lands. However, its stringent requirements are not responsive to today's environmental and economic realities. For the future of sustainable mineral extraction, there must be a better way. Adaptive management, a more flexible planning process, should be used on public lands to ensure greater leeway for operators, environmentalists, and local economies. By analyzing rural northeastern Utah's Uinta Basin's history and existing public land use plans, this Note applies adaptive management to the area to show how thinking outside the box can solve seemingly unsolvable problems.

* J.D. Candidate, BYU Law School, April 2024. Thanks to Professor John Fee and Professor Craig Galli for their input. Big thanks to the BYU Law Review editors and their helpful edits. And biggest thanks to Isaac for his patience always.

CONTENTS

INTRODUCTION

How we deal with climate change is (literally) one of the hottest topics in our contentious era. As a global citizen aware of threats created by man-made climate change, I support ending our reliance

on fossil fuels. But my upbringing in northeastern Utah's Uinta[1] Basin oil fields, where friends, neighbors, and family depend on mineral extraction to make ends meet, complicates the issue for me.

No matter the lens through which we view climate change and energy extraction, however, environmentalists and industrialists agree on one thing: federal and state land use planning, leasing, and permitting for energy extraction are slow, unresponsive processes wrapped in red tape. This red tape, the National Environmental Policy Act (NEPA), hinders progress, confuses the planning and permitting process, and ineffectively divides regulation between states, federal agencies, and operators.

Using the Uinta Basin, my often-overlooked homeland, as an example, I argue that the existing natural resource planning process is too rigid, discounts economic and environmental realities, and ultimately fails to protect the environment or the people. But there is a better way: Adaptive management, which provides more flexible and reactive mechanisms, is one answer to achieving a better balance between environmental protection and economic strength. Adaptive management would be most effective without NEPA, but it can also be integrated with existing frameworks if land use planners and managers think and act for a better future.

Part I outlines Uinta Basin land ownership from the beginning of the United States and discusses local realities today. Part II discusses the Basin's existing land use plans and planning mechanisms. Part III addresses problems with current land use and resource management plans. Part IV introduces adaptive management, and Part V applies an adaptive management model to the Basin's land use plans. Ultimately, I present a more agreeable land use planning process that is more responsive to our rapidly changing world.

1. This name is derived from the Ute word "yoov-we-teuh" meaning pine forest or pine tree, and the Ute word has two transliterations. "Uinta" is used for natural features, like the Uinta Mountains or the Uinta Basin, while "Uintah" is used for names, like Uintah County (but to make it more confusing, there is an Uinta County, Wyoming.) *See* Forest Serv., *A Century of Stewardship: Early History of the Uinta National Forest*, U.S. DEP'T OF AGRIC., https://www.fs.usda.gov/detail/uwcnf/learning/history-culture/?cid=stelprdb5052885 (last visited Oct. 30, 2023).

I. UINTA BASIN BACKGROUND

A. Uinta Basin History

To understand the Basin's unique mix of land ownership, it's important to understand its recent settlement history, fraught with federal, indigenous, and local conflicts from the beginning.

The Basin's settlement begins with Utah's settlement. In 1848, the United States took most of the western United States from Mexico in the Treaty of Guadalupe Hidalgo.[2] After the United States gained control, Mormon settlers, who had arrived in the Salt Lake Valley in 1847, petitioned Congress for their own state which stretched from Mexico to Oregon.[3] Amid pre-Civil War dynamics of balancing slave and free states, and with an intent to limit Mormon influence, Congress instead split its new acquisition into two large territories: Utah Territory to the north and New Mexico Territory to the south.[4]

Despite Congress's denial, Mormon[5] leader Brigham Young continued to send scouts and settlers across the West to build a religious empire. Northeastern Utah's Uinta Basin was one conquest, but in 1861, a damning scouting report stopped Mormon Basin expansion before it even began. The report opined that the area was "one vast 'contiguity of waste'" and "measurably valueless, excepting for nomadic purposes, hunting grounds for Indians and to hold the world together."[6] Later settlers agreed with the report's assessment. One newcomer reportedly burst into tears on her first view of her new Basin home.[7] Her despair is understandable—the Basin can certainly be inhospitable. A mile-high, bowl-shaped depression on the Colorado Plateau in northeastern Utah ringed by the Wasatch Mountain Range on the west, the Uinta Range on the

2. Glen M. Leonard, *The Mormon Boundary Debate Question in the 1849–50 Statehood Debates*, 18 J. OF MORMON HIST. 114, 114 (1992).

3. *Id.* at 118–21.

4. *Id.* at 132–33.

5. Following academic custom, I will refer to historical members of The Church of Jesus Christ of Latter-day Saints as "Mormon." I mean no disrespect by this appellation.

6. Jedediah S. Rogers. "*One Vast 'Contiguity of Waste'": Documents from an Early Attempt to Expand the Mormon Kingdom into the Uinta Basin, 1861*, 73 UTAH HIST. Q. 249, 250 (2005).

7. Settler Mary Brown, who came from Heber, Utah, in the 1880s, stated, "When we came around that mountain pass and looked into this valley of sagebrush and rabbitbrush, Oh, how I cried." LAMOND TULLIS, A SEARCH FOR PLACE: EIGHT GENERATIONS OF HENRYS AND THE SETTLEMENT OF UTAH'S UINTAH BASIN 161–62 (2010).

north, and the Tavaputs Plateau (the Book Cliffs) to the south, the Basin has little appeal on its arid surface.[8]

But despite its apparent barrenness, the Basin proved valuable to the federal government, both as a buffer against Mormon influence and as a place to forcibly settle native peoples. In 1861, President Abraham Lincoln carved out almost the entire area for the Uintah Valley Reservation by executive order.[9] The Reservation's creation, ostensibly to solve the "Indian problem," forced central Utah tribes to the Basin, where they had seasonally roamed but never permanently settled.[10] Twenty years later in 1882, the federal government created a second reservation, the Uncompahgre (modern-day Ouray) Reservation to force the Colorado Ute bands' settlement following uprisings and massacres in Meeker.[11] In 1886, the two reservations were merged as the Uinta-Ouray Ute Indian Reservation, with administrative activities centered in Fort Duchesne.[12]

Even though nearly four million acres were set aside for indigenous residents, Reservation lands were disrespected from the beginning. Newcomers, including miners, outlaws, homesteaders, and herders, breached treaties and trickled into the Reservation from other corners of Utah and the West, claiming the land for themselves.[13] In 1880, enough outsiders had moved into the area that the Utah Territorial Legislature created Uintah County on February 8, 1880, which ignored Reservation boundaries and stretched from Wasatch County to the Colorado border.[14]

Throughout the late 19th Century, congressional actions continued to shrink the Reservation. In 1888, when prospectors discovered gilsonite on Reservation land, miners tricked tribal leaders into signing a treaty, later ratified by Congress, that

8. DORIS K. BURTON, A HISTORY OF UINTAH COUNTY: SCRATCHING THE SURFACE, 3–4 (1998).

9. *See* MAP OF UINTAH VALLEY RESERVE, FORMED BY ORDER OF PRESIDENT ABRAHAM LINCOLN (1861), https://collections.lib.utah.edu/details?id=358892 (last visited Oct. 18, 2023).

10. Clifford Duncan, *The Northern Utes of Utah*, *in* HISTORY OF UTAH'S AMERICAN INDIANS, 167, 185, 189–93 (Forrest S. Cuch ed., 2000).

11. *Id.* at 195–96

12. *Id.* at 195.

13. BURTON, *supra* note 8, at 8.

14. *Id*.

annexed a 7,040-acre strip from the Reservation.[15] That same year, Congress also passed the Dawes Act, which allotted plots to individual tribe members to supposedly protect Indigenous peoples' property rights.[16] Under the Dawes Act, individual tribal members were given title to land within the Reservation, which the government held in trust until a certain time period, and then the land was reverted fee simple to the individual tribal member.[17] But instead of giving Native people a better life, this policy meant most Native people in the Basin ended up selling their plots and losing their property rights.[18]

At the turn of the century, with the Indian allotment program established and the surviving indigenous peoples with their own land claims, President Roosevelt and Congress formally opened the remaining unallotted reservation land to white homesteaders in 1905.[19] This settlement decision allowed people (my ancestors included) to claim native land for $1.25 per acre, or about $40 per acre today.[20] Any unclaimed land remained in federal hands, managed by the General Land Office, the precursor to today's Bureau of Land Management (BLM).[21] After opening the reservation, only a quarter of the reservation remained in Native hands, leading to boundary disputes that continue today.[22] Other federal agencies also claimed their share of President Lincoln's

15. John Barton, *Uinta Basin*, *in* FROM THE GROUND UP: THE HISTORY OF MINING IN UTAH, 378, 387–88 (Colleen Whitley ed., 2006).

16. *Dawes Act*, NAT'L ARCHIVES (Feb. 8, 2022), https://www.archives.gov/milestone-documents/dawes-act.

17. *Native American Ownership and Governance of Natural Resources*, NAT. RES. REVENUE DATA, https://revenuedata.doi.gov/how-revenue-works/native-american-ownership-governance (last visited Oct. 18, 2023).

18. *See* Casey McClellan, To Hold the World Together: A Uinta Basin Homesteading History, 1905–1930, (Mar. 18, 2021) (undergraduate honors thesis, Brigham Young University) (on file at https://scholarsarchive.byu.edu/studentpub_uht/183). The Dawes Act ended in 1934 when Congress passed the Indian Reorganization Act, which gave remaining Native lands to the federal government to be held in trust. *Id.*

19. PROCLAMATION 581—OPENING THE UINTA INDIAN RESERVATION LANDS, UTAH (1905), https://www.presidency.ucsb.edu/node/206410.

20. *Id.*

21. *Id.*

22. *See* Hagen v. Utah, 510 U.S. 399 (10th Cir. 1994) (starting the infamous Ute Tribe v. Utah line of cases in the Tenth Circuit, which all consist of battles over criminal law sovereignty in the former and current reservation boundaries).

envisioned Reservation lands, including what became the Ashley National Forest[23] and Dinosaur National Monument.[24]

The Basin's current land ownership and administration reflects its unusual settlement history: 59.4% of Uintah County,[25] 44.8% percent of Duchesne County,[26] and almost ninety percent of Daggett County[27] are owned and managed by the BLM, the National Forest Service, the Bureau of Reclamation, and the National Parks Service, with 15.34% managed by the Ute Tribe in Uintah and Duchesne Counties.[28] State Institutional Trust Land Association-owned (SITLA) land checkerboards across federal land, totaling 6.85% of the Tri-County area (Uintah, Duchesne, and Daggett Counties).[29] The remaining land is privately owned or managed by other state agencies.[30] Below is a current map of Uinta Basin land ownership:

23. Charles DeMoisy, Jr., *Early History of Ashley National Forest*, U.S. FOREST SERV. 1, 1–6, https://www.fs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb5421800.pdf.

24. *Dinosaur National Monument History and Culture*, NAT'L PARK SERV., https://www.nps.gov/dino/learn/historyculture/index.htm (last visited Oct. 18, 2023).

25. UINTAH COUNTY GENERAL PLAN 2 (2021), https://cms1files.revize.com/uintahcountyut/document_center/CommunityDevelopment/General%20Plan%202017%20Update%202-1-2021.pdf.

26. DUCHESNE COUNTY GENERAL PLAN 27 (2018), https://www.duchesne.utah.gov/wp-content/uploads/2018/10/2018-General-Plan-Amdt-IRAs.pdf.

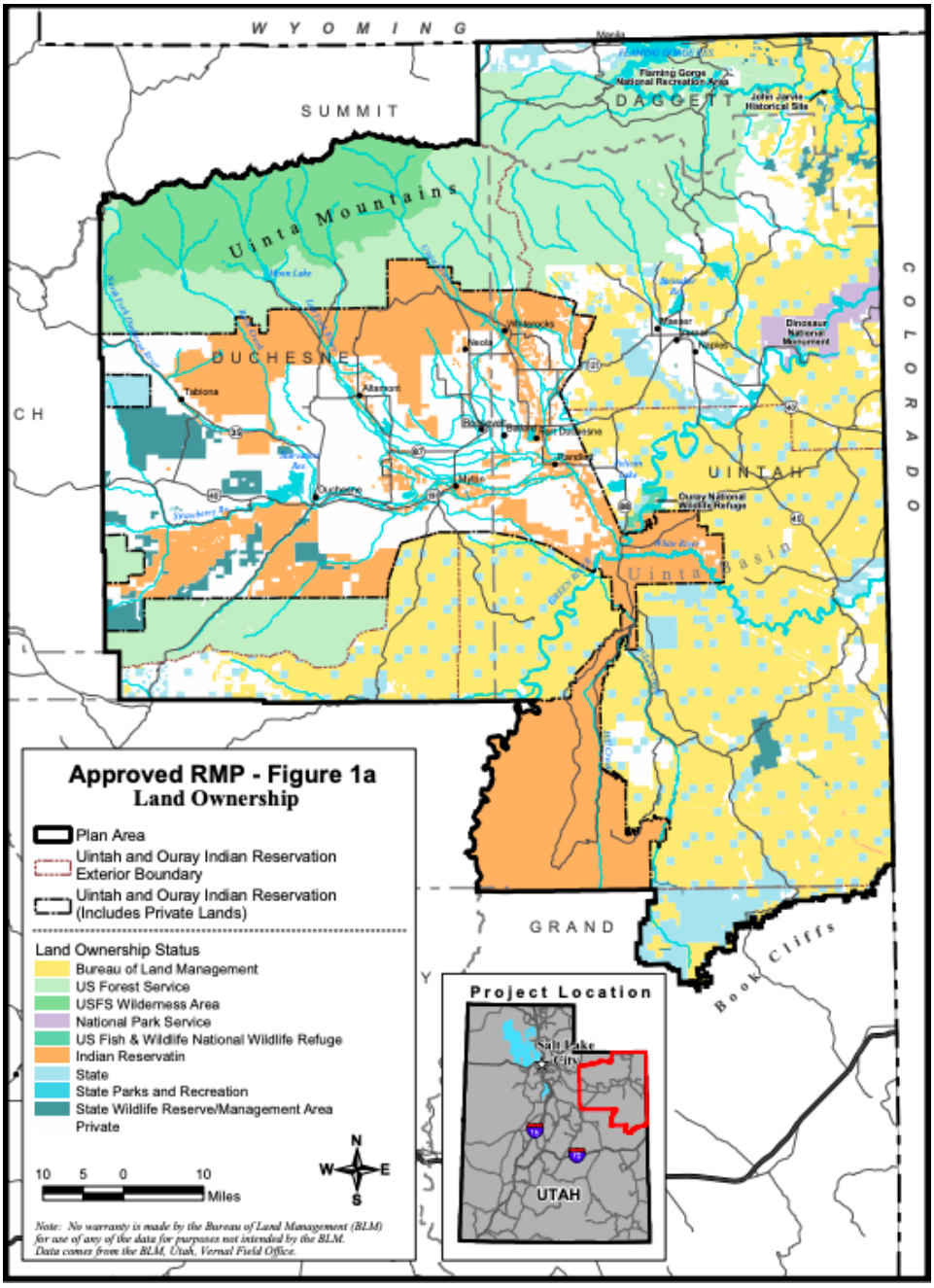
27. DAGGETT COUNTY GENERAL PLAN 10 (2008), https://www.daggettcounty.org/DocumentCenter/View/7848/Ordinance-17-17-Daggett-Chap-8-RMP-FINAL?bidId=.

28. VERNAL AREA RESOURCE MANAGEMENT PLAN 4 (2008), https://eplanning.blm.gov/public_projects/lup/68145/86218/103392/VernalFinalPlan.pdf [hereinafter VERNAL RMP].

29. *Id.*

30. *Id.*

Figure 1: Uinta Basin Land Ownership[31]



B. The Basin Today

At first glance, the Basin may have no purpose but "to hold the world together,"[32] but it is what's below the surface that puts it on the map. Basin residents today rely heavily on oil and gas extraction and are frustrated by perceived federal downscaling of drilling permits.[33]

The Uinta Basin's oil and natural gas output is Utah's biggest. In 2022, Duchesne County produced 29,184,533 barrels of oil, and Uintah County produced 11,560,505.[34] For comparison, in 2022 the next-highest producing county, San Juan County, produced 3,031,336 barrels.[35] The Basin's 2023 output is on track with 2022 levels, with Duchesne County reporting 19,287,200 barrels and Uintah County already reporting 6,697,557 as of July 2023.[36] Natural gas extraction is similarly high. In 2022, Duchesne County produced 55,803,583 McF,[37] and Uintah County Produced 155,185,628 McF.[38] In 2022, Duchesne County produced 55,803,583 McF,[39] and Uintah County Produced 155,185,628 McF. [40] As of July of 2023, Duchesne County has already extracted 32,889,444 McF and Uintah County extracted 97,513,110 McF.[41]

The revenue from all oil and gas in the state was $4 billion in 2018, and oil and gas drilling and supporting industries like refineries, transportation, and maintenance made up a total of 2.5%

31. VERNAL RMP, *supra* note 28, at 201 fig. 1a.

32. Rogers, *supra* note 6, at 250.

33. This perception is, however, not reality: the Biden administration continues to lease and permit at comparable numbers to the Trump administration. *See* Taylor McKinnon, *Biden Administration Oil, Gas Drilling Approvals Outpaces Trump's*, CTR. FOR BIOLOGICAL DIVERSITY (Jan. 24, 2023), https://biologicaldiversity.org/w/news/press-releases/biden-administration-oil-gas-drilling-approvals-outpace-trumps-2023-01-24.

34. *Oil Production by County (Past 5 Years)*, UTAH DEP'T OF NAT. RES. https://oilgas.utah.gov/oilProdVolPerCountyPerYear.php (last visited Oct. 20, 2023).

35. *Id.*

36. *Id.*

37. One McF is equal to one cubic foot.

38. *Natural Gas Production by County (Past 5 Years)*, UTAH DEP'T OF NAT. RES., https://oilgas.utah.gov/gasProdVolPerCountyPerYear.php (last visited Oct. 20, 2023).

39. *Id.*

40. *Id.*

41. *Id.*

of Utah's gross state product.[42] And the more drilling allowed on public lands, the better for Utah economically: half of oil and gas revenues from federal lands belongs to the state.[43]

The economic benefits are not only important to governments—they also matter to individual Basin residents. The oil and gas fields provide high-paying jobs, and those who work in them perceive public lands as a vital part of their individual and family livelihoods. For example, in Duchesne County in 2021, the average energy extraction job paid $73,877,[44] 46% higher than the County's average wage of $50,581.[45] Uintah County energy worker salaries exhibit similar advantages with a $74,113 average,[46] 66% higher than the county-wide average of $44,583.[47] This higher wage makes oil field jobs desirable for a population that has few postsecondary education opportunities.[48]

A 2008 study that accompanied the most recent BLM Vernal Planning Area Resource Management Plan indicates just how important activities on public lands are to Basin residents. The study found that 21.8% of households in Uintah County relied *directly* on permitted activities on public lands, with 11.1% of Duchesne County respondents reporting the same.[49] Additionally, in Uintah County, 83% of respondents felt public land resource uses were "very important" to the overall quality of life in their community, and in Duchesne County, 81.2% of respondents agreed.[50] 55.3% of Uintah County residents and 45.4% of Duchesne

42. Michael D. Vanden Berg, *Utah's Energy Landscape*, 127 UTAH GEOLOGICAL SURV. 1, 8–9 (2020), https://energy.utah.gov/wp-content/uploads/Utahs-Energy-Landscape-5th-Edition.pdf.

43. 30 U.S.C. § 191(a)–(b).

44. U.S. Bureau of Lab. Stat., *Employment and Wages Data Viewer*, U.S. DEP'T OF LAB., https://data.bls.gov/cew/apps/table_maker/v4/table_maker.htm#type=18&from=2020&to=2021&qtr=1&own=5&ind=10&area=49047&supp=1 (last modified Sept. 7, 2022).

45. *Id.*

46. *Id.*

47. *Id.*

48. As of 2021, only 17.3% of people had a bachelor's degree or higher in Uintah County. *QuickFacts-Uintah County, Utah*, U.S. CENSUS BUREAU, https://www.census.gov/quickfacts/uintahcountyutah (last visited Oct. 14, 2023). Similarly, only 14.5% in Duchesne County had a bachelor's degree or higher. *Id.*

49. U.S. BUREAU OF LAND MGMT., VERNAL RES. MGMT. PLAN app. M at M-2 (2008), https://eplanning.blm.gov/public_projects/lup/68145/86506/103676/M_-_Utah_Public_Lands_Study-Key_Social_Survey_Findings_for_Daggett,_Duchesne_and.pdf.

50. *Id.* at M-4.

County respondents wanted to increase exploration for oil and gas development on public lands.[51] Finally, 29.4% of all Tri-County respondents reported that they or other members of their households participate in oil and gas exploration and development, compared to the next highest county block in the same category: only 6.7% of respondents in Grand and San Juan Counties reported similar reliance on fossil fuel extraction.[52] Thus, more than in other parts of Utah, oil and gas are economically and culturally important to Basin residents.

Despite the economic incentives, however, there are shortcomings to oil and gas extraction that even locals recognize: nitrous oxide (NOx) and other chemicals released by drilling create dangerous air quality conditions, including ground level ozone.[53] Even low amounts of ozone affect breathing, and the Basin's winter ground ozone levels are concerning.[54] With ozone present, people with asthma and other respiratory illnesses are at higher risk, and children are especially harmed because their lungs are still developing.[55] And ozone and its effects negatively impact quality of life: a 2019 report by the Utah Department of health showed that adult asthma hospitalization rates in the Tri-County area nearly double the rest of Utah's average.[56]

Another greenhouse gas present because of oil and gas drilling is methane, which is a primary component of natural gas and is a potent greenhouse gas.[57] Methane leaks make up as much as three to five percent of the Basin's total energy output.[58] Although a

51. *Id.*

52. RICHARD S. KRANNICH, PUBLIC LANDS AND UTAH COMMUNITIES: A STATEWIDE SURVEY OF UTAH RESIDENTS 47 tbl.8 (2008).

53. *Ozone in the Uinta Basin: Ozone Basics*, UTAH DEP'T OF ENV'T QUALITY, https://deq.utah.gov/air-quality/ozone-in-the-uinta-basin (last updated Oct. 20, 2023).

54. *Id.*

55. *Ground-Level Ozone Pollution: Health Effects*, U.S. ENV'T PROT. AGENCY, https://www.epa.gov/ground-level-ozone-pollution/health-effects-ozone-pollution (last updated May 24, 2023).

56. Utah Dep't of Health and Hum. Serv., *Health Indicator Report of Asthma Hospitalizations*, PUB. HEALTH INDICATOR BASED INFO. SYS. (May 24, 2023), https://ibis.health.utah.gov/ibisph-view/indicator/view/AsthHosp.LHD.html.

57. John C. Lin, Ryan Bares, Benjamin Fasoli, Maria Garcia, Erik Crosman & Seth Lyman, *Declining Methane Emissions and Steady, High Leakage Rates Observed Over Multiple Years in a Western US Oil/Gas Production Basin*, SCI. REPS., Nov. 2021, at 1, 2, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8595340.

58. *Id.*

recent study[59] from Utah State University found that the Basin oil and gas industry has cut methane emissions by half since 2013, unfavorable economic conditions and less drilling, rather than better regulatory measures, caused the decrease in emissions.[60] If leakage regulation improved, not only would the climate be better off, so would oil and gas companies.

In summary, the Uinta Basin is a unique place that relies on mineral extraction to survive. And its complex historical background, leading to multiple landowners, managers, and issues, complicates the way the federal government and operators plan resource management and land use.

II. THE UINTA BASIN PLANS

This Part discusses land use plans used by the BLM, the Ute Tribe, SITLA, and Uintah & Duchesne Counties in the BLM's Vernal Planning Area, a 5,518,895 acre[61] area that stretches across the Basin. Although many factors, including permit issuance and market forces, contribute to drilling availability and emissions, this Note focuses on land use plans themselves, which designate where and how much oil and gas drilling is allowed.

A Federal Land Use and Resource Management Plans

1. BLM

The BLM manages the largest share of the Vernal Planning Area, controlling 30% of the 5,518,895-acre total.[62] Under guidance set forth in the Federal Land Policy and Management Act,[63] BLM land use planning focuses on "multiple use and sustained yield[,]"[64] meaning that different land uses in the same area are central to BLM's mission. To implement multiple and sustained use, the BLM periodically creates Resource Management Plans (RMPs). Per federal regulations,[65] Field Managers create each RMP with

59. *See generally id.*
60. *Id.*
61. VERNAL RMP, *supra* note 28, at 4.
62. *Id.*
63. 43 U.S.C. §§ 1701–87.
64. 43 U.S.C. §§ 1701(a)(7), 1732(a).
65. 43 C.F.R. § 1601.0-4 (2023).

planning guidance from the BLM Planning Handbook and undertake data collection related to natural resources in the area.[66] RMP creation is an arduous process and requires multiple drafts, public comment periods, and environmental evaluations.[67]

The BLM's Vernal Area RMP specifies all areas where drilling can occur and if there are any restrictions. In the Planning Area, 1.7 million acres of the 1.9 million acres of federal mineral estate are available for energy extraction and approved for leasing.[68] Once the RMP designates land as available, it remains available until a new RMP replaces it.[69] Updating the RMP (or any RMP) to respond to new information is near impossible: the Vernal RMP can only be amended or revised "if *major* changes are needed or to consider a proposal or action that is not in conformance with the plan."[70] Even amendments require "the appropriate level of environmental analysis."[71] For more minor changes, like spelling corrections, the RMP can be "maintained" without a formal process, but that maintenance is "limited to refining, documenting, and/or clarifying previously approved decisions."[72]

The RMP is only evaluated every five years unless "unexpected actions, new in-formation [sic], or significant changes in other plans, legislation, or litigation trigger[] an [early] evaluation."[73] Since its creation in 2008, there have been twenty-five RMP maintenance actions and no amendments.[74] Additionally, one evaluation was published in 2022 which used data collected in 2014 (even though evaluations should happen every five years, budget and staff restraints have probably contributed to the limited evaluations). The evaluation hinted to more responsive changes using adaptive

66. U.S. BUREAU OF LAND MGMT., LAND USE PLANNING HANDBOOK H-1601-1 27 (2005), https://www.ntc.blm.gov/krc/uploads/360/4_BLM%20Planning%20Handbook%20H-1601-1.pdf.

67. *Id.*

68. VERNAL RMP, *supra* note 28, at 61.

69. *Id.*

70. *Id.* at 62 (emphasis added).

71. *Id.*

72. *Id.*

73. *Id.* at 63.

74. U.S. BUREAU OF LAND MGMT., VERNAL RMP FIVE-YEAR EVALUATION REPORT 4 (2014), https://eplanning.blm.gov/public_projects/68145/200138838/20058799/250064981/VFO%20RMP%20Five%20Year%20Evaluation%202014.pdf.

management, a process that will be discussed in Parts III, IV, and V, but still only allows for "minor modifications or adjustments . . . without amendment or revision of the plan as long as assumptions and impacts disclosed in the analysis remain valid and broad-scale goals and objectives are not changed."[75]

The RMP also ignores air quality and effects of permitted activities on air quality. This is because the "BLM does not have regulatory control over air quality issues, either on public lands or on Tribal or state lands. BLM relies on the agency with jurisdiction over air quality to set regulatory standards and criteria to protect the air quality in a particular area."[76] Similarly, the RMP establishes "goals," but does not require any type of aggregated air quality assessment.[77]

Overall, the Vernal Area RMP is the most comprehensive planning document for public lands in the Basin, but it lacks important functions and is not updated or reviewed as often as it should be.

2. Tribal Plans

The Ute Tribe manages 846,669 acres, 15.34% of the Vernal Planning Area.[78] Land use planning on tribal lands is complicated thanks to multiple land holders and various interests. Individual Tribe members might hold their own land fee simple[79] and may lease it if they have a permit, but all remaining land is retained by the tribe in trust and managed by the federal government. Within remaining Tribe-owned land, various title holders, including the Ute Indian Tribe itself, multiple landowners such as the Ute Indian Allotted Land, Ute Indian Tribe, and Ute Distribution Corporation Jointly Managed Indian Trust Minerals, make their own planning

75. *Id.* at 3.
76. VERNAL RMP, *supra* note 28, at 26.
77. *Id.*
78. *Id.*
79. *See supra* Part I.

decisions.[80] The Tribe also receives administrative and technical support from the Bureau of Indian Affairs (BIA) and the BLM.[81]

Although the BIA and Ute Tribe plan where drilling is allowed, the BLM retains ultimate leasing authority[82] and must account for "the best interest of the Indian mineral owner . . . including, but not limited to: economic considerations . . . ; probable financial effect on the Indian mineral owner; leasability of land concerned; need for change in the terms of the existing lease; marketability; and potential environmental, social, and cultural effects."[83] Often, this "best interest" is an economic one.

B. State and Local Land Use Plans

1. SITLA and State Plans

SITLA manages 6.85% of the Vernal Planning Area.[84] Like in other western states, when public land surveys began, four of the thirty-six sections in each federally organized township were set aside for public beneficiaries like schools and civic buildings; all revenues from SITLA land goes to these beneficiaries.[85] SITLA parcels are sometimes exchanged with federal land holdings; recently Utah lawmakers and federal land managers swapped SITLA parcels that were within the newly expanded Bears Ears National Monument for BLM land outside the Monument.[86] SITLA's planning and leasing requirements are set forth in the Utah Administrative Code, which allow for leasing wherever the State Division of Oil, Gas, and

80. *Uintah and Ouray Reservation*, BUREAU OF INDIAN AFFS., https://web.archive.org/web/20170516195623/https:/www.bia.gov/cs/groups/xieed/documents/document/idc1-022549.pdf.

81. Bureau of Indian Affs., *Mineral Leasing on Individual Indian and Tribal Lands*, U.S. DEP'T OF THE INTERIOR., https://www.bia.gov/service/leasing/mineral-leasing (last visited Oct. 12, 2023).

82. 43 C.F.R. § 3161.1 (2023).

83. 25 C.F.R. § 211.3 (2023).

84. *Id.*

85. *Beneficiaries*, TR. LANDS ADMIN., https://trustlands.utah.gov/our-impact/beneficiaries (last visited Oct. 12, 2023).

86. Kyle Dunphey, *Why Utah is Trading over 160,000 Acres of Bears Ears with the Federal Government*, DESERET NEWS (May 18, 2022, 5:27 PM) https://www.deseret.com/utah/2022/5/18/23124492/bear-ears-land-transfer-approved-utah-lawmakers-goes-to-us-congress-trump-biden-national-monument.

Mining allows.[87] There are no planning restrictions as long as an operator receives a permit from the Division.[88]

2. County Plans

In the Vernal Planning Area, 22.17% (1,223,791 acres) is privately owned.[89] Utah law allows oil and gas extraction on private land if an applicant receives a leasing permit from the State Division of Oil, Gas, and Mining[90] and if they comply with local ordinances.

The Uintah County General Plan and RMP, two separate planning documents, are united in purpose: drill, baby, drill.[91] The plans hardly discuss any environmental concerns or mitigation strategies and broadly allow for drilling in most zones.[92] Although the County recognizes the importance of a diverse economy,[93] it does little to address these concerns—most of the future land development plans in western Uintah County remain open to mineral extraction.[94] To defend this single-mindedness, the RMP cites economic success: in 2014, oil production in Utah was valued at $3.2 billion, and of the five largest oil-producing fields in Utah, four were in the Uintah Basin.[95] For local resource management planners, the "least restrictive stipulations" are the best.[96]

Duchesne County also emphasizes that local access to "public lands for all forms of energy development" is vital.[97] Duchesne County asks federal land managers to expedite leases, increase lease approval, and improve access to minerals,[98] stating that "[a]ll available, recoverable solid, fluid and gaseous mineral resources in the subject lands should be seriously considered for contribution or

87. Utah Admin. Code R850-21 (2019).

88. *Id.*

89. VERNAL RMP, *supra* note 28, at 4.

90. UTAH CODE ANN. 40-6-9.5 (West 2023).

91. UINTAH COUNTY GENERAL PLAN, *supra* note 25, at 36; UINTAH COUNTY RESOURCE MANAGEMENT PLAN 16–19, https://cms1files.revize.com/uintahcountyut/Res%2008-12-2019%20R1-%20Uintah%20Resource%20Management%20Plan-%20CC%20Approved.pdf.

92. *See generally id.*

93. UINTAH COUNTY GENERAL PLAN, *supra* note 25, at 36.

94. *Id.* at 21.

95. UINTAH COUNTY RESOURCE MANAGEMENT PLAN, *supra* note 91, at 16 ("The energy industry is vital to the Uintah County economy.").

96. *Id.* at 17.

97. DUCHESNE COUNTY GENERAL PLAN, *supra* note 26, at 66.

98. *Id.* at 69–70.

potential contribution to the state's economy and the economies of the respective counties."[99]

Ultimately, though, the counties have little land to lease and can only issue policy recommendations to the federal government. They rely heavily on the BLM to plan for local needs, but it's clear from the local plans' language that they would approve drilling projects without many restrictions if they had the choice.

III. THE PROBLEM WITH RESOURCE MANAGEMENT PLANS

This Part critiques the Basin plans and the RMP creation process in general. As the world's climate ventures into unknown territory, NEPA's time-consuming process fails to address the realities of our increasingly unpredictable environment. Additionally, NEPA washes its hands of pollution and air quality, forcing other agencies to deal with consequences of the BLM's (sometimes) poor planning.

A. The World is Changing, and NEPA Can't Keep Up

As global temperatures steadily climb toward a 1.5-degree Celsius increase, we inch closer to tipping points, which are temperatures and circumstances that humans have never experienced.[100] Once a tipping point is crossed, we have reached uncharted territory and cannot predict environmental effects and outcomes. But despite climate change's urgency, our current planning mechanisms, including the NEPA process which guides the creation of RMPs, remain hindered by tradition and caution.[101]

NEPA requires all major federal actions to be preceded by comprehensive environmental review of the action's effects, first by an Environmental Assessment (EA), and then, if the EA predicts significant environmental impacts, a more involved Environmental Impact Statement (EIS).[102] EAs and EISs require significant time

99. *Id.* at 54.

100. Courtney Lindwall, *Climate Tipping Points Are Closer than Once Thought*, NAT'L RES. DEF. COUNCIL (Nov. 15, 2022), https://www.nrdc.org/stories/climate-tipping-points-are-closer-once-thought.

101. As early as 1997, the Council on Environmental Quality recognized that the NEPA process fails to account for "[c]hanges in conditions—whether as a result of surprises from nature or human action . . . ." *See The National Environmental Policy Act: A Study of Its Effectiveness After Twenty-five Years*, COUNCIL ON ENV'T QUALITY 32 (Jan. 1997) https://ceq.doe.gov/docs/ceq-publications/nepa25fn.

102. 40 C.F.R. § 1501.3 (2023).

(the average review takes 3.4 years),[103] money, and expertise. Through the EA and EIS requirement, NEPA forces federal land managers, when creating RMPs, to "stop and think"[104] about a plan's potential impacts *before* a plan's creation.

Although NEPA's cautious approach prevents potentially harmful actions and has become more efficient over time,[105] there is still room for improvement.[106] NEPA's front-loaded decision-making means that any meaningful changes to a plan count as a "major federal action,"[107] which requires additional environmental reviews. Although creating and changing EAs does not require as much time, analysis, or resources as an EIS, the time it takes to react to unpredictable environmental impacts prevents NEPA-permitted plans from being as responsive as they need to be. And revisions, or more major changes, only apply when a plan becomes outdated or obsolete.[108]

The Vernal Area RMP, like most other RMPs, shares these problems. Any meaningful changes require substantial reviews, making it difficult to react to the changing realities in the Basin and beyond.

B. Disconnect Between Land Managers and Air Quality Monitors

A second problem with the Vernal Planning Area RMPs (and RMPs in general) is the disconnect between agencies who manage activities that cause emissions. The Vernal Area RMP requires the BLM to comply with "all local, state, federal, and tribal air quality regulations[,]" but does not coordinate with the monitoring agencies, plan for any contingencies, or set any goals itself—it simply delegates air quality monitoring to the EPA and the Utah Division of Air Quality, taking little responsibility for its decision-making.[109]

103. Piet deWitt & Carole A. deWitt, *How Long Does It Take to Prepare an Environmental Impact Statement?*, 10 ENV'T PRAC. 164–74, (Dec. 5, 2008), https://www.cambridge.org/core/journals/environmental-practice/article/abs/research-article-how-long-does-it-take-to-prepare-an-environmental-impact-statement/C1B14ECB03EBB159A2CE6B3A43CB5FAB.

104. Becker v. Fed. R.R. Admin., 999 F. Supp 240, 251 (D. Conn. 1996).

105. *See The National Environmental Policy Act: A Study of Its Effectiveness After Twenty-five Years*, COUNCIL ON ENV'T QUALITY (discussing how NEPA's concurrent study schedule, rather than consecutive, improved approval and planning time periods).

106. Mark Squillance, *Rethinking Public Land Use Planning*, 43 HARV. ENV'T L. REV. 415 (2019).

107. New Mexico *ex rel.* Richardson v. Bureau of Land Mgmt., 565 F.3d 683, 684 (10th Cir. 2009).

108. VERNAL RMP, *supra* note 28, at 62.

109. *Id.* at 70.

The lack of oversight is a problem. The Basin's air quality is notoriously bad, due in large part to oil and gas wells which release NOx and volatile organic compounds into the air, which combine with sunlight to create dangerous gases like ozone. During winter inversions, ozone levels sometimes exceed the EPA attainment level of 70 ppb.[110] If the Basin continues its non-attainment longer than three years in a row, the EPA will establish a new federal air quality plan—a State Implementation Plan (SIP) required by the Clean Air Act[111]—until attainment is reached.[112] The BLM's planning mechanisms directly impact how much land is used for drilling; the more land that is available, the more leases are approved, and the more air pollution there will be.

Instead of heaping on yet another requirement to operators and land managers, more flexible and holistic planning processes are a better option. This Note's final two Parts discuss a land use planning process that does just that: adaptive management.

IV. ADAPTIVE MANAGEMENT

Adaptive management has been around since the 1970s[113] and is, in theory, part of the NEPA process. In practice, however, it is not widely used outside of certain ecosystems. Fully implementing adaptive management requires a paradigm shift[114] and maybe means an end to the NEPA environmental review process as we know it. But even if only certain aspects of adaptive management are applied, the planning process can become a more responsive and effective way to acknowledge stakeholder concerns while also protecting the environment for the future.

110. *Fact Sheet; Air Quality in the Uintah Basin*, UTAH STATE UNIV., BINGHAM ENTREPRENEURSHIP & ENERGY RSCH. CTR., (July 2018), https://www.usu.edu/binghamresearch/files/2-pagehandoutUBairquality.pdf.

111. 42 U.S.C. § 7401.

112. *Ozone in the Uinta Basin*, *supra* note 53.

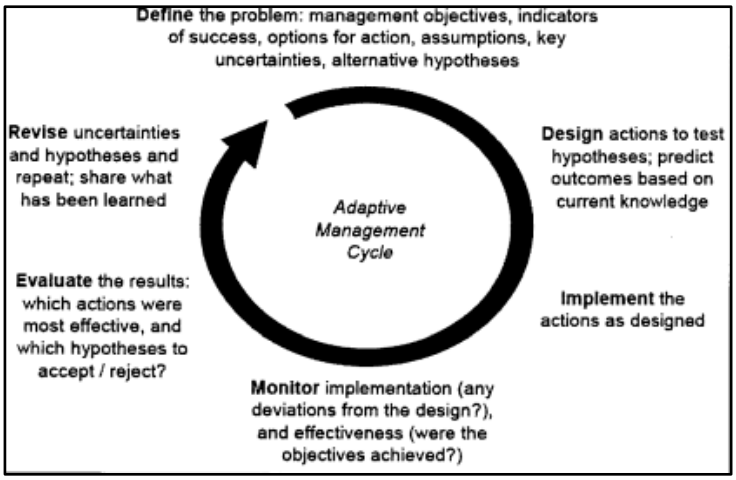
113. ADAPTIVE ENVIRONMENTAL ASSESSMENT AND MANAGEMENT (C.S. Holling ed., 1978) (creating the academic discussion for adaptive management in environmental law).

114. *See* Sam Kalen, *Public Land Management's Future Place: Envisioning a Paradigm Shift*, 82 MD. L. REV. 240 (2023); Mark Squillance, *Rethinking Public Land Use Planning*, 43 HARV. ENV'T. L. REV. 415 (2019).

A. Adaptive Management in Land Use Policy

Adaptive management has no standardized definition, but put simply, it is "learning by doing."[115] Instead of NEPA's process, which mandates that all land use decisions be made before a plan is adopted, adaptive management continually tracks planning decisions' effects because "[l]earning in adaptive management occurs through the practice of management itself, with adjustments as understanding improves."[116]

Figure 2: Adaptive Management Cycle[117]



Adaptive management is intended to be a multi-step, flexible cycle, with continual monitoring and actions that reflect scientific realities.[118] This "learning by doing" approach rejects the precautionary principle, which resists any unknown outcomes, and instead urges state actors to do something—anything—before it is too late.

115. Holly Doremus, *Precaution, Science, and Learning While Doing in Natural Resource Management*, 82 WASH. L. REV. 547, 550 (2007).

116. BYRON K. WILLIAMS & ELEANOR D. BROWN, ADAPTIVE MANAGEMENT: THE U.S. DEPARTMENT OF THE INTERIOR APPLICATIONS GUIDE 5 (2012), https://www.doi.gov/sites/doi.gov/files/uploads/DOI-Adaptive-Management-Applications-Guide-WebOptimized.pdf.

117. Martin Z. P. Olszynski, *Failed Experiments: An Empirical Assessment of Adaptive Management in Alberta's Energy Resources Sector*, 50 U.B.C. L. REV. 697, 708 (2017).

118. Robin Kundis Craig & J.B. Ruhl, *Designing Administrative Law for Adaptive Management*, 67 VAND. L. REV. 1, 1 (2014).

Although adaptive management plans are tailored to area-specific needs, adaptive management plans generally retain a flexible, monitoring-based pattern. The following adaptive management plan examples illustrate how they work.

Central California's Lewiston Dam, finished in the early 1960s, damaged the Trinity River's ecosystem when it slowed the river's flow and blocked fish runs.[119] To remedy this harm, the Bureau of Reclamation adopted an adaptive management plan to restore and protect the river's ecology in 2000. Within the Trinity River Restoration Program, multiple working groups monitor different aspects of the river's health: fish, sediment deposition, water temperature, streamflow, etc. The groups also meet frequently to share data and plan next steps, shifting timelines and adjusting the plan when one change affects another part of the river.[120] This program is a success story for adaptive management implementation.

Another adaptive management example, although less successful, is one of the first ever implemented: the Glen Canyon Dam Adaptive Management Plan. Like the Trinity River plan, it employs multiple working groups to improve downstream resources while delivering water quotas to the parched Southwest. But the program had a major flaw: although managers experimented with different water flow levels, there were no action requirements and thus, no permanent system improvement or solutions.[121] Because of this confusion about how to implement changes, Glen Canyon Dam Plan has earned some criticism,[122] but the project remains a step in the right direction and has been a helpful pilot project, hopefully paving the way for more successful plans.

Adaptive management has also been implemented with various degrees of success in multiple National Forest RMPs, the Missouri River, and the Everglades.[123]

119. *The Trinity River*, TRINITY RIVER RESTORATION PROGRAM, https://www.trrp.net/program-structure/background (last visited Oct. 18, 2023).

120. *Science and Adaptive Management*, TRINITY RIVER RESTORATION PROGRAM, https://www.trrp.net/restoration/adaptive-management (last visited Oct. 18, 2023).

121. Lawrence Susskind, Alejandro E. Camocho & Todd Schenk, *A Critical Assessment of Collaborative Adaptive Management in Practice*, 49 J. APPLIED ECOLOGY 47, 48 (2011), https://besjournals.onlinelibrary.wiley.com/doi/10.1111/j.1365-2664.2011.02070.x.

122. *Id.*

123. ADAPTIVE MANAGEMENT FOR ECOSYSTEM RESTORATION: ANALYSIS AND ISSUES FOR CONGRESS, CONG. RSCH. SERV., (2011), https://www.everycrsreport.com/reports/R41671.html.

B. Obstacles to Broadly Implementing Adaptive Management

Although adaptive management is more responsive to change than current planning processes, there are some obstacles to wider implementation, including a lack of experience with adaptive management plans, political mistrust, and NEPA's strict framework.

Because adaptive management is underutilized, there is often a disconnect between a plan's aspirations and its actual application by inexperienced managers. Even in places like Canada, where adaptive management is more widely adopted, there are still gaps between policy and practice. In a comprehensive survey of adaptive management plans and climate change mitigation in Alberta's oil fields, Canadian environmental law professor Martin Olszynski found that adaptive management, although written into many plans, was applied inconsistently and proved unenforceable because of a dearth of experience.[124] This is a circular problem, but one that cannot be addressed without managers' willingness to try adaptive management and work through the unknowns.

Another obstacle facing adaptive management is political pressure, leading to unclear expectations for stakeholders. Although the U.S. Department of the Interior has issued adaptive management planning guidance and advisement since 2007,[125] Congress quashed broader-scale attempts of implementing adaptive management into BLM handbooks in 2016.[126] Recognizing this resistance, the BLM has attempted to introduce adaptive management language in more veiled terms; in November 2022, the BLM released an Instruction Memorandum that hinted at an adaptive approach that could change oil and gas leasing based on "changing circumstances, updated policies, and new information . . . ."[127] Despite the BLM's support, current guidelines combined with a lack of political support prevent more robust implementation of adaptive

124. Olszynski, *supra* note 117, at 794.

125. *Adaptive Management*, U.S. DEP'T OF THE INTERIOR (2009), https://www.doi.gov/sites/doi.gov/files/uploads/TechGuide-WebOptimized-2.pdf.

126. For an in-depth discussion of this process and the amended language, see Robert L. Glicksman & Jarryd Page, *Adaptive Management and NEPA: How to Reconcile Predictive Assessment in the Face of Uncertainty with Natural Resource Management Flexibility and Success*, 46 HARV. ENV'T L. REV. 121 (2022).

127. *Instruction Memorandum 2023-101: Oil and Gas Leasing – Land Use Planning and Lease Parcel Reviews*, BUREAU OF LAND MANAGEMENT (Nov. 21, 2022), https://www.blm.gov/policy/im-2023-010.

management. Successful adaptive management plans require specific goals and implementation strategies, as well as a willingness to act despite uncertainty.[128] Without broader consensus and clearer guidance, however, adaptive management remains on the fringe.

Adaptive management also plays an uncomfortable dance with NEPA's strict requirements. Courts and other actors, now accustomed to NEPA's "hard look[,]"[129] are wary of adaptive management's more flexible approach and are uncomfortable with "scientific uncertainty."[130] Some courts have struck down adaptive management plans as arbitrary and capricious under the Administrative Procedure Act when written in terms too flexible to satisfy NEPA's requirements.[131] Courts have also rejected adaptive management plans that do not have enough potential alternatives.[132] The more flexible a plan, the better the plan can adapt to new environmental problems. But enabling this flexibility would require requires major changes to NEPA, which is an unlikely prospect.[133]

C. Adaptive Management Best Practices

Although adaptive management would likely be at its most successful without NEPA, it can still be implemented within the bounds of existing constraints as long as an adaptive management plan takes a "hard look" at mitigation measures and potential

128. *See* Mary Jane Angelo, *Stumbling Toward Success: A Story of Adaptive Law and Ecological Resilience*, 87 NEB. L. REV. 950, 991–1007 (2009) (discussing a Florida adaptive management plan that, despite some unintended consequences, reached its goals by leaning into uncertainty, accepting the sometimes-negative outcomes, and rolling with the punches).

129. Nat. Res. Def. Council, Inc. v. Morton, 458 F.2d 827, 838 (D.C. Cir. 1972).

130. Greater Yellowstone Coal., Inc. v. Servheen, 665 F.3d 1015, 1029 (9th Cir. 2011) ("[I]t is not enough to invoke 'adaptive management' as an answer to scientific uncertainty"). This quote, of course, demonstrates the qualms judges have with scientific analysis in general and why implementing adaptive management is so difficult.

131. *See* Klamath Siskiyou Wildlands Ctr. v. Boody, 468 F.3d 549, 557 (9th Cir. 2006) (striking down a BLM plan that allowed the agency to change the plan without a formal process when it vaguely "'contemplate[d]' a wide swath of changes"); W. Org. of Res. Councils v. U.S. Bureau of Land Mgmt., 2018, No. 4:20-cv-00076-GF-BMM, WL 3082475, at *7 (D. Mont. Aug. 3, 2022) (ruling a Power River Basin RMP as "arbitrary and capricious" under the APA because the RMP lacked viable alternatives and did not acknowledge down-stream effects of oil and gas projects).

132. *See* Save Our Cabinets v. U.S. Dep't of Agric., 254 F. Supp. 3d 1241, 1270–71 (D. Mont. 2017).

133. *See* Eric Biber, *Adaptive Management and the Future of Environmental Law*, 46 AKRON L. REV. 933 (2013) (addressing the limits of adaptive management).

outcomes. According to recent court decisions, taking a "hard look" requires more than "[a] mere listing of mitigation measures . . . ."[134] Uncertainty might be allowed, but vagueness is not.[135] Instead, a viable plan must have enforceable standards that trigger clear, adaptive responses, and a clear structure within the plan itself.[136] The Forest Service, which has applied adaptive management more broadly than any other agency, has learned to list potential actions in enough detail to pass a NEPA challenge.[137] And according to scholarship by Courtney Shultz and Martine Nie, as long as a plan is detailed enough and has enough predicted ranges of possible outcomes, "[c]ourts do not always require additional NEPA analysis when new information comes to light, as long as any changes in action and predicted effects are within the range of what was analyzed in the original NEPA document."[138]

Recent scholarship and reports have identified a few best practices for NEPA approval and successful adaptive management plan implementation. In a 2003 White House Council on Environmental Quality (CEQ) report, the CEQ identified specific factors like clear monitoring goals, baseline data collection, scientific and technical expertise to identify changes, and resources to monitor and respond to changes that would update the NEPA process and make it more responsive.[139] Expanding on these factors, Glicksman and Page have found a better balance between

134. Neighbors of Cuddy Mountain v. U.S. Forest Serv., 137 F.3d 1372, 1380 (9th Cir. 1998).

135. *See* High Sierra Hikers Ass'n v. Weingardt, 521 F. Supp. 2d 1065, 1083–84 (N.D. Cal. 2007).

136. *See* J.B. Ruhl & Robert L. Fischman, *Adaptive Management in the Courts*, 95 MINN. L. REV. 424, 464–66 (2010).

137. When creating an adaptive management plan, a Forest Service EIS

> must clearly identify the adjustment(s) that may be made when monitoring during project implementation indicates that the action is not having its intended effect or is causing unintended and undesirable effects. The EIS must disclose not only the effect of the proposed action or alternative but also the effect of the adjustment. Such proposal or alternative must also describe the monitoring that would take place to inform the responsible official during implementation whether the action is having its intended effect.

36 C.F.R. § 220.5(e)(2) (2020).

138. Courtney Schultz & Martin Nie, *Decision-Making Triggers, Adaptive Management, and Natural Resources Law and Planning*, 52 NAT. RES. J. 443, 458 (2012).

139. THE NEPA TASK FORCE, REPORT TO THE COUNCIL ON ENVIRONMENTAL QUALITY: MODERNIZING NEPA IMPLEMENTATION 50 (Sept. 2003), https://ceq.doe.gov/docs/ceq-publications/report/finalreport.pdf.

planning outright and planning with changes in mind.[140] These factors include (i) procedural strategies that identify and include stakeholders in order to improve the chance of buy-in and support, as well as substantive plans that include (ii) clear and specific management goals, (iii) accurate baseline data, (iv) identification of triggers that will require further action, (v) monitoring to determine when triggers are exceeded, and (vi) commitment to mitigation or predetermined ranges of possibilities if a trigger is exceeded.[141]

1. Stakeholder Inclusion

First, stakeholders, including the public, tribal governments, environmental advocates, cooperating agencies, and industry stakeholders, must support the adaptive management plan.[142] Stakeholder and public support not only satisfies NEPA[143] and other CEQ requirements,[144] but it also allows for multiple voices to express concerns and be heard. Indeed, "meaningful stakeholder participation should serve a central role in the management of natural resources and the regulation of land use and . . . regulatory processes should account for the uncertainty inherent in regulatory decisions by making such processes more adaptive."[145] Stakeholder inclusion also ensures that, because everyone has a voice, no dissatisfied party will be able to claim that it was not included. The downsides of stakeholder inclusion are that it takes time and it can be difficult for different groups with different needs to agree on a plan. But stakeholders don't all have to agree on everything—they only need their voices acknowledged, and at the end of the day, the land manager will make the decision.

2. Clear Management Goals

Next, plans themselves must prioritize clarity, include a range of possibilities, and create clear triggers that will allow for

140. Glicksman & Page, *supra* note 126, at 188–195.

141. *Id.*

142. *Id.* at 186.

143. 40 C.F.R. § 1501.7 (2023).

144. 40 C.F.R. § 1506.6(a) (2023).

145. Alejandro E. Camacho, *Beyond Conjecture: Learning About Ecosystem Management from the Glen Canyon Dam Experiment*, 8 NEV. L.J. 942, 943 (2008).

additional action within the limits of NEPA while avoiding the cumbersome EIS and EA processes.

One of the downfalls of the Glen Canyon Adaptive Management Plan was its lack of clear goals and "priorities."[146] To summarize UC Berkely's Eric Biber, adaptive management plans require straightforward goals to guide the plan, determine the costs and benefits of each decision, know what concessions stakeholders are willing to make, and know how to evaluate the plan's success.[147] The clearer the initial goals and the more transparent the plan, the more "indicat[ion] to the public and courts that the agency is not merely making a 'plan to make a plan' but has laid out (and considered the impacts of pursuing) a meaningful agenda in advance."[148] Although there will be some competing goals because of the agency's collaboration with multiple stakeholders, the agency's ultimate decision-making responsibility will be to set goals that make the most sense for the area based on input from all sides. Thoughtful goals will help stakeholders and the land manager maintain a vision of the bigger picture as well as guide the rest of the plan's implementation.

3. Baseline Data Identification

Without baseline data, there will be nothing to compare later measurements with.[149] And the data will ensure more accurate monitoring later in the process. Recording retrievable oil and gas levels, potential extraction methods, existing air quality levels, current extraction rates, and other metrics is vital for future goal setting and planning.

Such data gathering does not have to be from scratch, either. For example, existing data sets from scientists, operators, the United States Geospatial Survey, and other organizations are

146. *See* Sandra Zellmer & Lance Gunderson, *Why Resilience May Not Always Be a Good Thing: Lessons in Ecosystem Restoration from Glen Canyon and the Everglades*, 87 NEB. L. REV. 893, 930 (2009) ("The primary impediment to making the most of the opportunities created by experimentation on the Grand Canyon is Congress's unwillingness to articulate clear ecological priorities among conflicting societal values.").

147. Biber, *supra* note 133, at 955.

148. Glicksman & Page, *supra* note 126, at 189 (quoting Hells Canyon Pres. Council v. Connaughton, No. 3:11-cv-00023, 2012 WL 13047991, at *11 (D. Or. Aug. 10, 2012)).

149. *See generally* Holly Doremus, *Adaptive Management as an Information Problem*, 89 N.C. L. REV. 1455, 1475 (2011) (discussing how, when a system is in crisis, "it is too late to go back and generate historic data").

accessible and available if planners seek them out. And baseline conditions and data are already required for RMP creation and amendment. Augmenting existing data, completing more thorough investigations, and, when needed, performing new data collection, will better prepare planners and plans for inevitable future changes.

4. Trigger Identification

Next is knowing what data levels, or triggers, will require action in an adaptive management plan. Triggers are data points that will force land managers to act. One of the problems with the Glen Canyon Adaptive Management Plan was a lack of clear triggers—it "[did] not mandate when information gleaned from such experiments must be used to adjust the management protocols."[150] Knowing this, original and articulated triggers within a plan are vital for a plan's response to be legitimate. If a trigger is vague or non-existent, critics will undoubtably find that that adaptive management is used to avoid better planning and decision-making.[151] Additionally, "[t]riggers can be used to force adaptation in response to monitoring results [and] to provide underlying guarantees that important resources will be protected from serious, irreversible impacts from adaptive management experiments."[152] If they are specific, these triggers can also avoid the political back-and-forth that other decisions often create, thus reducing NEPA challenges in court.

5. Monitoring and Data Review Plan

As mentioned above, CEQ recognizes the importance of monitoring in NEPA processes. Indeed, without monitoring, "there can be no improved understanding of conditions or responses to management actions, and therefore, no informed adjustment of on-the-ground practices."[153] With the Glen Canyon Adaptive Management Plan, there were not enough technical experts to sift

150. Zellmer & Gunderson, *supra* note 150, at 930.

151. *See* HOLLY DOREMUS ET AL., MAKING GOOD USE OF ADAPTIVE MANAGEMENT 11 (2011), https://cpr-assets.s3.amazonaws.com/documents/Adaptive_Management_1104.pdf ("[A]daptive management can become a tool to rationalize uncertainty or cover flaws in initial decisions, rather than a mechanism for improving management over time.").

152. Biber, *supra* note 133, at 960.

153. Schultz & Nie, *supra* note 138, at 447.

through the data, thus making the monitoring a drain on resources instead of a meaningful and reactive tool.[154] The more focused monitoring, the better chance of recognizing trigger points that prompt policy change. In addition to consistent monitoring and recording, data must also be regularly analyzed—data collection does no good if the results are not applied.[155]

6. Action Commitment

Finally, adaptive management is meant to spur changes, so actions directed from the process require "rigorous implementation"[156] and commitment. The action commitment is probably the most difficult part of adaptive management to integrate with NEPA because years-long environmental reviews are required for any NEPA-controlled changes to a plan. But if the agency has described (in detail) possible actions and outcome ranges in the plan, that flexibility will likely still be NEPA compliant without another EA or EIS.[157] Indeed, courts have upheld adaptive management plans that described "fixed mitigation measures" even though the details of those measures were not yet set in stone.[158]

In summary, adaptive management plans are possible within the existing NEPA framework as long as the plans spell out mitigation possibilities in detail. With some foresight and willingness from all stakeholders to accept a little uncertainty, adaptive management plans would be an important innovation for planning in general.

V. ADAPTIVE MANAGEMENT IN THE UINTA BASIN

Using the best practices listed above, this final Part will suggest adaptive management applications for Uinta Basin land management

154. *See* Susskind et al., *supra* note 121, at 48.

155. *See* DOREMUS ET AL., *supra* note 151, at 12 ("Data must not sit on a shelf. The learning effort must include systematic and ongoing data interpretation and evaluation, as well as data sharing within and between agencies so that learning diffuses from one action to others.").

156. J.B. Ruhl, *Taking Adaptive Management Seriously: A Case Study of the Endangered Species Act*, 52 U. KAN. L. REV. 1249, 1278 (2004).

157. *See* Glicksman & Page, *supra* note 126, at 194.

158. *See* Nat'l Parks Conservation Ass'n v. Jewell, 965 F. Supp. 2d 67, 75–77 (D.D.C. 2013) (Where "general mitigation measures" that did not list the content of the mitigation plans were not arbitrary and capricious because although mitigation plans were not fleshed out in detail, the process for submitting mitigation plans was set forth in detail in the plan, meeting the NEPA "hard look" requirement).

plans to improve air quality and ensure a better quality of life for people in the Basin.

A. Stakeholder Inclusion

One of the biggest challenges of implementing adaptive management in the Uinta Basin would be the profound ideological differences between locals concerned with their livelihoods[159] and other groups who see the entire oil and gas industry as damaging.[160] These different viewpoints are at odds with the more collaborative approach required by adaptive management, but through specific strategies, a plan can build trust and give all sides the chance for their voice to be heard.

Holding meetings for *all* interested parties (not just the coordinating ones) and reaching out to different groups while in the planning phase would be useful ways to expand the planning process. The existing RMP already lists the local counties, as well as the Ute Indian Tribe and the State of Utah, as cooperating agencies in its creation,[161] but hearing from other interested parties like environmental groups, industry representatives, and scientific teams can shape the plan in a meaningful way that covers all bases and allows all people to be heard.

And the more groups involved, the more factions will be able to say that agencies heard their perspectives and increase support for a plan. Allowing for broad stakeholder representation will also guide goals and show each side that there are other things to consider than their own agendas.

B. Clear Management Goals

Next, the plan's content must be specific, focused, and flexible. Easing into adaptive management with a goal that appeals to all camps can help guide the adaptive management plan. As mentioned

159. *See generally* UINTAH COUNTY GENERAL PLAN, *supra* note 25, at 2 ("The energy resources in Uintah County create the primary industry in the region."); UINTAH COUNTY RESOURCE MANAGEMENT PLAN, *supra* note 91, at 16 ("The energy industry is vital to the Uintah County economy."); DUCHESNE COUNTY GENERAL PLAN, *supra* note 26, at 30 ("It is important to the County economy that public lands be properly managed for . . . energy production [and] mineral extraction . . . .").

160. *See Dirty Fuels*, S. UTAH WILDERNESS ALL., https://suwa.org/issues/dirty-fuels (last visited Sept. 19, 2023).

161. VERNAL RMP, *supra* note 28, at 56.

in the Utah State Study, 3–5% of methane produced in the Basin escapes into the atmosphere[162] along with NOx and other particulates. One specific goal to alleviate this escape would be to lower methane produced by energy extraction to 1–2%. Another goal along the same vein would be to keep all NOx emissions low enough to reach the EPA's current ozone attainment level of 70 ppb. These specific, measurable data points are not only better for the environment, but they are also better for operators because they keep as much energy protected as possible and help drilling companies operate more efficiently—a win-win.

Once this goal is met, the sky is the limit for what other goals can be implemented within the adaptive management plan. But new goals would require some type of mechanism or specific structure within the plan itself to show how to add new goals once previous goals are reached. These layers of planning might make an adaptive management plan more complicated, but if the mechanism for changing and adding goals is within the plan itself, the foundation is set for adaptive management to expand to other areas that are more typical in the adaptive management world—such as riparian health, protecting endangered species, and other environmental impacts—while still balancing oil and gas production.

C. Baseline Data

Baseline data is one of the most easily applicable aspects of adaptive management in the Basin—it's already done as part of the NEPA process.[163] The 2008 RMP featured a comprehensive survey from a few years before of oil and gas yet to be drilled,[164] and the United States Geological Survey also published a survey in 2010 estimating the total number of barrels still in the ground.[165] This baseline data can help operators know where they can be most efficient. The other vital data set is much more up-to-date: researchers at Utah State University and the Utah Division of Air

162. Lin et al., *supra* note 57, at 2.

163. U.S DEP'T OF THE INTERIOR BUREAU OF LAND MGMT., H-1601-1, LAND USE PLAN. HANDBOOK app. d at 2–3 (2005).

164. U.S DEP'T OF THE INTERIOR BUREAU OF LAND MGMT., MINERAL REPORT FOR THE VERNAL PLANNING AREA, (2002), https://eplanning.blm.gov/public_projects/lup/68145/86675/103847/MPR_100702.pdf.

165. U.S. GEOLOGICAL SURV., DDS-69-BB, OIL SHALE RESOURCES OF THE UINTA BASIN, UTAH, AND COLORADO (2010), https://pubs.er.usgs.gov/publication/dds69BB.

Quality annually track emissions and gas particulates in a comprehensive and ongoing survey.[166] There are four monitoring stations within the Basin that track levels of fine particulates, NOx, and methane.[167] These stations can also tell where certain particulates and gases are coming from, making the monitoring more targeted for improving the air quality. The existing data would probably be enough to satisfy this baseline requirement.

D. Trigger Identification

Next, knowing what conditions will cause mandatory action by operators and land managers will ensure transparency and better understanding of plans. Creating specific triggers that stakeholders can agree on will ensure greater success of a potential adaptive management plan. Stakeholders, especially locals and industry groups, would likely not accept a broad trigger, such as global warming degrees. Other, more tangible metrics are more realistic and would be more applicable to the Basin. One potential trigger could be the EPA's ppb ozone threshold, which is set to trigger a Uinta Basin State Action Plan in case of non-attainment.[168] Incorporating this existing level into an adaptive management plan will help all stakeholders know which metrics matter and what they must do to prevent mitigation actions that may be more drastic.

Additionally, identifying triggers will bring peace of mind to operators, who know that if they stay below certain levels, they will be able to keep drilling. And knowing that there will be penalties for operators who exceed the plan's limits will bring environmentalists on board as well. Although oil and gas drilling are used to fuel other similarly polluting industries, working with current and local solutions can improve Basin residents' health and hopefully help the industry to be more environmentally conscious.

166. *See* Lin et al., *supra* note 57, at 1.

167. *See generally* Ozone Overview and Standard Ozone SIP, UTAH DEP'T OF QUALITY, https://deq.utah.gov/air-quality/ozone-overview-and-standard-moderate-area-ozone-sip (last visited Oct. 18. 2023).

168. *Id.*

E. Monitoring Schema

Currently, evaluations are supposed to be monitoring data, but those evaluations happen only once every five years.[169] Having more frequent evaluation periods would be an important step in applying adaptive management. As mentioned, data from Utah State University and the Utah Division of Air Quality's existing air monitoring sites could determine when triggers are exceeded.[170] The United States Geological Survey also already monitors oil and gas levels in the ground.[171] As long as scientists and managers evaluate and apply this data, then there is less need for bigger, more cumbersome planning processes like creating entirely new RMPs.

Some industry stakeholders may also want monitoring for economic conditions, but that goes beyond an RMP's control. RMPs and adaptive plans focus on the direct effects of the drilling activity, not the economic ones. Economic and market information could be useful for creating an adaptive management plan and could be part of industry representation concerns at the beginning of the process, but it would be difficult to integrate into triggers or thresholds. Instead, throughout the management process, industry groups that are part of the plan's working group or team can present this data to land managers. The BLM could use the data to consider its next move, but it does not need to be a part of a land use plan because it is not a land use-specific metric.

F. Action Commitment

An action commitment from stakeholders and managers is likely the hardest pill to swallow in an adaptive management plan. Hopefully, with improved communication and representation at the early planning stages, more operators will be willing to conform to the action commitment. Their conformity will, of course, depend on what the triggers are, but stakeholders must understand the need for action. Reducing fossil fuel emissions can be as simple as plugging up non-producing wells and making pipelines more efficient, and then seeing if air quality improves. If it helps, then the

169. VERNAL RMP, *supra* note 28, at 50.

170. *See* Lin et al., *supra* note 57, at 2.

171. *See* U.S. GEOLOGICAL SURV., *supra* note 165.

plan has succeeded, and other goals and monitoring can replace the existing ones.

But if emissions remain high, other solutions may also be required, including limiting leasing in certain areas or initiating some type of regional cap-and-trade system. Not all stakeholders will agree with these more stringent measures, but a tiered system of patching up leaks, monitoring, and waiting to see the results would allow companies to continue drilling without waiting for an environmental assessment that could be much harsher in its requirements. Overall, adaptive management would bring better balance to the Basin.

CONCLUSION

The Uinta Basin is a unique corner of the world, and maybe it does "hold the world together."[172] Allowing its people to continue their livelihoods while mitigating the negative effects of its largest revenue source keeps the Basin growing while also working toward a cleaner and more sustainable future. As our world changes, flexible and thoughtful solutions—not rote and stale processes—will save us. With thoughtful changes and an established mitigation strategy, adaptive management is one strategy the Basin (and the planet) need to ensure sustainable extraction for years to come.

172. Rogers, *supra* note 6, at 250.

Byte a Carrot for Change: Uprooting Problems in Data Privacy Regulations

Sarah Terry[*]

There is a growing gap between technology advancement and a lagging regulatory system. This is particularly problematic in consumer data privacy regulating. Companies hold collected consumer data and determine its use largely without accountability. As a result, ethical questions that carry society-shaping impact are answered in-house, under the influence of groupthink, and are withheld from anyone else weighing in.

This Note poses a solution that would address multiple data privacy regulation issues. Namely, an incentive approach would help even out the information-imbalanced system. Incentives are used as tools throughout intellectual property law to foster commercial progress, discourage trade secrets, and protect consumers. These goals can also be achieved through integrating an incentive into consumer data privacy regulating.

This Note first highlights major issues in the current consumer data privacy federal regulatory landscape. Next, this Note proposes and outlines a narrow FTC whistleblower incentive, unearthing how an incentive would alleviate each major regulatory issue. Finally, this Note discusses eight compelling reasons for the incentive, and ultimately confronts and rebuts its drawbacks.

* J.D., BYU Law School, 2023. I thank Professor Clark Asay for his thoughtful insight and guidance on this Note during the drafting process. I also thank the fabulous BYU Law Review team of editors who contributed their time and many talents. Finally, I thank my family for supporting me in so many meaningful ways throughout this endeavor.

CONTENTS

INTRODUCTION

Shortly after the launch of ChatGPT, the most disruptive technological innovation in decades, OpenAI Chief Technology Officer (CTO) Mira Murati and OpenAI co-founder Elon Musk each publicly called for regulators to get involved.[1] These calls are rooted in a notion both have openly articulated: the consumer data collected and used by ChatGPT will shape society—for better, for worse, and for now—in largely unknown ways.[2] Unknown, that is, by the public and by regulators. Companies, on the other hand, may already have a clear picture of the impacts their consumer data strategies cause. However, companies are without obligation to disclose their discoveries—including known harms—and instead hold research and findings as proprietary trade secrets.[3] As a result, ethical questions about data usage are answered in-house, under the influence of groupthink, in the interest of commercialism, and are withheld from anyone else weighing in.[4]

The United States has largely taken a consent-based approach to data privacy, letting clicked boxes or continued use[5] perform "moral magic."[6] Such consent is called into question and seriously undermined when even the experts on artificial intelligence are not

1. *See* Steve Mollman, *ChatGPT Must Be Regulated and A.I. 'Can Be Used by Bad Actors,' Warns OpenAI's CTO*, FORTUNE (Feb. 5, 2023, 1:45 PM), https://fortune.com/2023/02/05/artificial-intelligence-must-be-regulated-chatgpt-openai-cto-mira-murati; Jyoti Narayan, Krystal Hu, Martin Coulter & Supantha Mukherjee, *Elon Musk and Others Urge AI Pause, Citing 'Risks to Society'*, REUTERS (Apr. 5, 2023, 6:22 AM), https://www.reuters.com/technology/musk-experts-urge-pause-training-ai-systems-that-can-outperform-gpt-4-2023-03-29. *But see* David Shepardson, *Tesla Recalls 362,000 U.S. Vehicles over Full Self-Driving Software*, REUTERS (Feb. 16, 2023, 5:34 PM), https://www.reuters.com/business/autos-transportation/tesla-recalls-362000-us-vehicles-over-full-self-driving-software-2023-02-16 (discussing one instance Musk resisted regulators and disagreed with oversight findings for the AI-based autopilot feature in Tesla cars).

2. *See* Narayan et al., *supra* note 1.

3. *See* Lital Helman, *Trade Secrets and Personal Secrets*, 55 U. RICH. L. REV. 447, 447–63 (2021) (discussing the differences and overlap between the approaches to trade secret law and data privacy law, namely, where personal data can "belong" to both a company's trade secrets and an individual).

4. *See* Mollman, *supra* note 1.

5. *See* Daniel J. Solove, *Murky Consent: An Approach to the Fictions of Consent in Privacy Law*, 104 B.U. L. REV. (forthcoming 2024) (manuscript at 4), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4333743 [hereinafter Solove, *Murky Consent*] (discussing the fictions of consent and an argument for a revised "murky consent" approach).

6. *See* Heidi M. Hurd, *The Moral Magic of Consent*, 2 LEGAL THEORY 121, 121–46 (1996).

fully sure how consumer data is used.[7] Put simply, consumers might be aware their data is collected, but are largely unaware of the extent to which it is used and how such use is affecting them.

In 2021, a Facebook insider leaked papers demonstrating consumer data has already been used to target propagandistic untruths toward the most susceptible individuals—those most eager to believe and promulgate the sensational material.[8] A recent controversy involving Twitter in summer of 2022 revealed national security issues are at stake because an insider demonstrated that Twitter has more data than it realizes or knows what to do with.[9] But these companies did not come forward with this information willingly. These businesses decided transparency was outside their best interest. Instead, a small number of employees brought these issues to the attention of the public, risking their careers, reputations, financial security, and personal relationships, and facing legal liability against deep-pocketed tech giants.

Something needs to change. But after more than a decade of effort, federal lawmakers have yet to pass bipartisan bills that cohesively regulate consumer data privacy.[10] The American consuming public has called for it, and now even data lobbyists are on board.[11] Congress has struggled to craft these laws in part because of the difficulty in striking the right balance between consumer and commercial interests. Specifically, an overarching policy seeks federal data regulations that both adequately address

7. *See* Solove, *Murky Consent*, *supra* note 5, manuscript at 2.

8. *See* Dipayan Ghosh & Ben Scott, *Facebook's New Controversy Shows How Easily Online Political Ads Can Manipulate You*, TIME MAG. (Mar. 19, 2018, 12:38 PM), https://time.com/5197255/facebook-cambridge-analytica-donald-trump-ads-data.

9. *Transcript: Twitter Whistleblower Testimony to Senate Judiciary Committee*, TECH POL'Y PRESS (Sept. 13, 2022), https://techpolicy.press/transcript-twitter-whistleblower-testimony-to-senate-judiciary-committee [hereinafter *Transcript*].

10. *See* Press Release, Fin. Servs. Comm., Financial Services Committee Advances McHenry's Data Privacy Act (Feb. 28, 2023), https://financialservices.house.gov/news/documentsingle.aspx?DocumentID=408620 (discussing the latest development in federal action to enact data privacy law, the Data Privacy Act of 2023); *see also* Data Privacy Act of 2023, H.R. 1165, 118th Cong. (2023).

11. *See* Alfred Ng, *Privacy Bill Triggers Lobbying Surge by Data Brokers*, POLITICO (Aug. 28, 2022, 7:02 AM), https://www.politico.com/news/2022/08/28/privacy-bill-triggers-lobbying-surge-by-data-brokers-00052958 (demonstrating that data industry lobbyists have boosted their spending after introduction of the American Data Privacy and Protection Act).

harms that may result from consumer data collection, while also refraining from over-burdening commercial innovation.[12]

This Note proposes one "ingredient" to help solve the many challenges regulators face. Namely, a garden-variety, incentive-based "carrot" to help regulate and address defects in federal data privacy law. Incentives are an integral part of intellectual property law.[13] Incentives are put into place to protect creatives, inventors, and brand rights while still facilitating innovation, benefits to society, and judicial efficiency.[14] Patent law works in various ways to discourage trade secrets while still providing protection for original inventions through the use of incentives.[15] An incentive in data privacy law could facilitate a similar goal of (1) promoting commercial progress and discouraging total trade secrecy while also (2) remaining rooted in protecting consumers. Specifically, this Note argues for a narrowly drafted whistleblower incentive and protection law relating to consumer data privacy and "unfair" or "deceptive" data practices under Section 5 of the Federal Trade Commission Act (FTC Act).

This metaphorical "carrot" has eight particularly appetizing features: (1) it furthers both prongs of the above-stated policy in a balanced way; (2) it offers an efficient enforcement tool to proportionately address harms and risks of harms; (3) it serves as a counterweight balancing the highly secretive nature of data information flow in a lagging regulatory system; (4) it presents a palatable piece for the Federal Trade Commission (FTC) to build on and learn from as the FTC confront challenges in rulemaking; (5) it can be easily grafted into any current or future bundle of regulations; (6) it fosters transparency and accountability; (7) it encourages consistent investment by companies in following regulations; and (8) it offers a way to promote deterrence while providing self-sufficient funding for the FTC.

12. *See* Danielle Keats Citron & Daniel J. Solove, *Privacy Harms*, 102 B.U. L. REV. 793, 820 (2022) (discussing the goals of data privacy enforcement).

13. *See* Jeanne C. Fromer, *Expressive Incentives in Intellectual Property*, 98 VA. L. REV. 1745, 1746–89 (2012) (discussing the "broadening of incentive possibilities [in intellectual property law] from a utilitarian position").

14. *See id.*

15. *See* J. Jonas Anderson, *Secret Inventions*, 26 BERKELEY TECH. L.J. 917, 919 (2011) ("Patents are often conceptualized as a means of luring secret inventions out of the dark, shadowy cave of trade secrecy, and into the bright, public sunlight of the patent system.").

In Part I, this Note first gives an overview of the FTC—the primary avenue to currently address data privacy laws in the United States at the federal level. Part II gives a synopsis of some relevant deficiencies and inefficiencies that result from the current FTC approach. This highlights the need for the incentive and protection law because the "carrot" addresses each of the mentioned deficiencies in at least one important way.

Part III builds on one specific inadequacy discussed above to help illustrate why the incentive is a solution to a not-fully-unearthed problem. Namely, the current system inadequately addresses known harms and is unable to gain a better understanding of lesser-known harms. There is first a discussion of known harms that result from inadequate consumer data privacy laws. Next, an analogy illustrates a fundamental reason why adequately addressing harms in consumer data privacy is uniquely challenging.

Part IV outlines and explains each carrot characteristic and why each is necessary. Part V analyzes and expands on each of the eight bits of the carrot byte, and Part VI confronts each of the carrot's potential drawbacks and challenges.

I. A NARROW OVERVIEW OF THE FTC'S CURRENT DATA PRIVACY REGULATORY LANDSCAPE AND ITS PITFALLS.

For more than two decades, the FTC has been the nation's privacy agency, consistently at the forefront of the public debate on online privacy.[16] The FTC's mission statement is to protect "the public from deceptive or unfair business practices and from unfair methods of competition through law enforcement, advocacy, research, and education."[17] This mission is balanced with an additional goal: to avoid unduly burdening legitimate business activity by maintaining a "vibrant economy fueled by fair competition."[18]

16. *See* Trade Regulation Rule on Commercial Surveillance and Data Security, 87 Fed. Reg. 51273 (proposed Aug. 22, 2022) [hereinafter Trade Regulation Rule].

17. *About the FTC*, FED. TRADE COMM'N, https://www.ftc.gov/about-ftc (last visited Oct. 9, 2023).

18. *See id.*

The FTC Act, set forth in 15 U.S.C. §§ 41–58, as amended, empowers the FTC to:

- "prevent unfair methods of competition and unfair or deceptive acts or practices in or affecting commerce;"[19]
- "seek monetary redress and other relief for conduct injurious to consumers;"[20]
- through rulemaking authority, "prescribe rules defining with specificity acts or practices that are unfair or deceptive, and establishing requirements designed to prevent such acts or practices;"[21]
- "gather and compile information and conduct investigations relating to the organization, business, practices, and management of entities engaged in commerce;"[22]
- "make reports and legislative recommendations to Congress and the public[;]"[23] and
- act as "the only federal agency with both consumer protection and competition jurisdiction in broad sectors of the economy."[24]

Section 5 of the FTC Act provides the principal legal authority for the FTC to regulate privacy and data security.[25] Section 5 prohibits "deceptive" or "unfair" commercial acts or practices.[26] A representation, omission, or practice is deceptive under Section 5 if it "is likely to mislead consumers acting reasonably under the circumstances," and is material to consumers—that is, it would

19. *Federal Trade Commission Act*, FED. TRADE COMM'N, https://www.ftc.gov/legal-library/browse/statutes/federal-trade-commission-act (last visited Oct. 30, 2023).

20. *Id.*

21. *Id.*

22. *Id.*

23. *Id.*

24. *About the FTC*, *supra* note 17.

25. *See* 15 U.S.C. § 45; *see also A Brief Overview of the Federal Trade Commission's Investigative, Law Enforcement, and Rulemaking Authority*, FED. TRADE COMM'N, https://www.ftc.gov/about-ftc/mission/enforcement-authority (last updated May 2021). The FTC has rulemaking and enforcement authority over several federal laws, including the Children's Online Privacy Protection Act, 15 U.S.C. §§ 6501–6506; Telephone Consumer Protection Act (Do-Not-Call Provisions), 15 U.S.C. §§ 6151–6155; CAN-SPAM Act, 15 U.S.C. §§ 7701–7713; HITECH Act, 42 U.S.C. §§ 17937 and 17953; and Gramm-Leach-Bliley Act, 15 U.S.C. § 6801 et seq. However, this Note narrowly focuses on Section 5 of the FTC Act.

26. 15 U.S.C. § 45.

likely affect the consumer's conduct or decision with regard to a product or service.[27] Generally, an act or practice is unfair under Section 5 if (1) "it causes or is likely to cause substantial injury to consumers[,]" (2) consumers cannot reasonably avoid the injury themselves, and (3) benefits to consumers or competition do not outweigh the injury.[28]

Currently, the FTC approaches privacy and data security through case-by-case enforcement and general policy work.[29] FTC enforcement actions are based on allegations that certain practices violate Section 5 of the FTC Act or other federal or state statutes to the extent these actions have harmed or pose the risk of harm through physical security, cause reputational or economic injury, or involve undesirable intrusions into the daily lives of customers.[30]

The FTC has brought actions including:

- an action against Abika.com, for secretive collection and sales of detailed consumer phone records obtained through false pretenses and without required consumer consent;[31]
- an action against CafePress, for its alleged failure to put into place and apply reasonable measures to protect consumers' personal information, and subsequently covering up a resulting breach;[32]
- an action against Twitter seeking civil penalties, permanent injunction, and monetary relief for collecting consumers' phone numbers and email addresses to improve social media account security, but also deceptively using that data to allow companies to target advertising in violation of an existing settlement agreement;[33]
- an action against Google and its subsidiary YouTube for allegedly illegally collecting personal information from

27. Cliffdale Assocs., 103 F.T.C. 110, 174 (1984).

28. 15 U.S.C. § 45(n).

29. *See* Trade Regulation Rule, *supra* note 16.

30. *See id.* at 51278.

31. *See, e.g.*, Complaint for Injunctive and Other Equitable Relief, Fed. Trade Comm'n. v. Accusearch Inc., 570 F.3d 1187 (10th Cir. 2009) (No. 06CV0105D).

32. *See, e.g.*, Complaint, Residual Pumpkin Entity, LLC, Nos. C-4768, C-4769, 2022 WL 2355879 (F.T.C. June 23, 2022).

33. *See, e.g.*, Complaint for Civil Penalties, Permanent Injunction, Monetary Relief, and other Equitable Relief, United States v. Twitter, Inc., No. 3:22-cv-3070, 2022 WL 1768852 (N.D. Cal. May 25, 2022).

children without their parents' consent, where the FTC and New York Attorney General settled with the companies for a $170 million settlement;[34] and

- an action against Facebook for allegedly violating a 2012 FTC order by deceiving users about their ability to control the privacy of their personal information. This resulted in a historic penalty with a settlement that imposed $5 billion, and significant requirements to boost accountability and transparency.[35]

Overall, the FTC brings actions when companies fail to comply with representations made to consumers about their data privacy and security practices, or when companies fail to implement reasonable security measures to protect sensitive information.

Notably, there is no private right of action for unfair or deceptive trade practices at the federal level—consumers must rely on the FTC to bring action.[36] However, because many states have state-level equivalents, state attorney generals can help fill the gaps that the FTC misses.[37] Additionally, the FTC generally cannot seek civil monetary damages except when a respondent has violated a prior settlement agreement.[38] So instead, typical FTC enforcement remedies require prohibiting acts complained of, remediating problematic acts, deleting wrongfully obtained information, modifying privacy policies, establishing a comprehensive privacy program, performing biennial audits for twenty years, and record-keeping and reporting obligations.[39] The FTC can only seek monetary penalties once a company is legally committed to one of the aforementioned remedies, the company subsequently violates

34. *See, e.g.*, Complaint, Fed. Trade Comm'n. v. Google LLC, No. 1:19-cv-02642 (D.D.C. Sept. 10, 2019).

35. *See, e.g.*, Stipulated Order for Civil Penalty, Monetary Judgment, and Injunctive Relief, United States v. Facebook, Inc., No. 1:19-cv-02184 (D.D.C. Jul. 24, 2019).

36. *See* Allan Bruce Currie, *A Private Right of Action Under Section Five of the Federal Trade Commission Act*, 22 HASTINGS L.J. 1268, 1268 (1971); Moore v. N.Y. Cotton Exchange, 270 U.S. 593, 603 (1926).

37. *See* Press Release, Federal Trade Commission, Commission Seeks Public Comment on Collaboration with State Attorneys General (June 7, 2023).

38. To obtain civil monetary penalties under the FTC Act, generally, the FTC must first find a respondent has violated a previously entered cease-and-desist order. Once that hurdle is cleared, the FTC must then bring a subsequent enforcement action for a violation of that order. *See* 15 U.S.C. § 45(b)–(n).

39. *See* Daniel J. Solove & Woodrow Hartzog, *The FTC and the New Common Law of Privacy*, 114 COLUM. L. REV. 583, 606–25 (2014).

that agreement, the FTC becomes aware, and the FTC consequently takes action—again.[40]

There is, however, an exception where the FTC can seek monetary penalties in the first instance. Namely, the FTC Act authorizes the FTC to impose civil penalties for first-time violations of "duly promulgated trade regulation rules."[41] However, the FTC has yet to put into place trade regulation rules regarding data privacy.

The FTC published an advanced notice of proposed rulemaking (ANPR) seeking public comment on potential data privacy rules in August of 2022.[42] The ANPR noted that the FTC is "beginning to consider the potential need for rules and requirements regarding commercial surveillance and lax data security practices."[43] The ANPR includes a numbered list of ninety-five questions, with additional queries within many of the numbered questions.[44] The ANPR noted this is done to "generate a . . . record about prevalent commercial surveillance practices or lax data security practices that are unfair or deceptive, as well as about efficient, effective, adaptive regulatory responses" to help improve the FTC's enforcement work, and to inform Congress or other policymakers as they work toward reform.[45]

The size of some of the settlements and extensive obligations imposed on certain companies through FTC actions regarding data privacy might be thought to indicate an effective federal system is in place. However, there are some major deficiencies and barriers the current data privacy regulatory system must overcome. This Note's proposed "carrot" would help the FTC alleviate each of the noted barriers in at least one important way. The FTC's deficiencies are discussed below.

40. *See id.*
41. Trade Regulation Rule, *supra* note 16, at 51280; *see* 15 U.S.C. § 45(m).
42. *See* Trade Regulation Rule, *supra* note 16, at 51280.
43. *Id.* at 51277.
44. *Id.*
45. *Id.*

II. FACING FTC CHALLENGES

A. Inadequate Staffing Resources Face Increasingly Sophisticated Technology

As a preliminary matter, the FTC is without adequate resources to respond to the issues imposed by data privacy underenforcement. In September 2021, the FTC released a report to Congress, which revealed its Division of Privacy and Identity Protection had around only forty to forty-five employees.[46] With such limited resources and broad tasks—all while facing a giant industry—"short-staffed" seems like a less than adequate adjective.

To illustrate the disproportionate staff to prevalence problem, one study's statistics illustrate current commercial fraud extensiveness and its devastating potential due to inadequate enforcement: identity fraud involving use of consumers' personal information amounted to twenty-four billion dollars stolen last year, ensnaring fifteen million U.S. consumers.[47]

Further, the technology enabling these results continues to grow more sophisticated in its collection and analysis of data, thus potentially becoming more threatening. This is the case for ChatGPT, where there is already evidence it is being used to generate phishing scams to help scammers with imperfect English and poor grammar.[48] There are also forums for ChatGPT collaboration among fraudsters, demonstrating that ChatGPT is only beginning to be put to nefarious use.[49]

46. *See* FED. TRADE COMM'N, FTC REPORT TO CONGRESS ON PRIVACY AND SECURITY (2021), https://www.ftc.gov/system/files/documents/reports/ftc-report-congress-privacy-security/report_to_congress_on_privacy_and_data_security_2021.pdf.

47. *See* John Buzzard, *2022 Identity Fraud Study: The Virtual Battleground*, JAVELIN (Mar. 29, 2022), javelinstrategy.com/2022-Identity-fraud-scams-report.

48. *See OPWNI: Cybercriminals Starting to Use ChatGPT*, CHECK POINT RESEARCH (Jan. 6, 2023), https://research.checkpoint.com/2023/opwnai-cybercriminals-starting-to-use-chatgpt.

49. *Id.*

B. The FTC's Enforcement Is Inefficient Due to Difficulty in Detectability and Lagging Regulator Expertise

Even assuming adequate resources, the FTC's enforcement power is inefficient to a point of diminishing deterrence.[50] Such inefficiency creates inconsistency in compelling companies to make proper practices a priority. The inefficiency largely exists because of the process the FTC must go through before it can seek civil penalties. The first step is likely the most formidable hurdle the FTC faces: the FTC must almost miraculously become aware of largely undetectable data practices that may be unfair or misleading. This alone is a major barrier in achieving efficiency and deterrence.

Deceptive or unfair data practices are largely undetectable because of an imbalance of information between the data industry and government. "If 'information is the "lifeblood" of effective governance,' the current prospects for effectively governing tech look dim."[51] The United States government has taken a hands-off approach, valuing technological innovation with an optimism in companies' self-regulation.[52] As a result, companies use trade secrecy arguments "to insulate themselves from oversight."[53] This has led companies to hold almost all the cards as they confront and determine how to resolve critical questions that "directly affect[] human behavior, individual rights, and freedom."[54]

Even if the first hurdle is cleared and the FTC becomes aware of a potential issue and decides to investigate, the FTC's investigative process can also be inefficient in discovering problems.[55] Inefficiency aside, the FTC has also failed to detect when a company skirted direct questions about compliance with crafty answers. An example of this was articulated in congressional testimony from Peiter Zatko, Twitter's former head of security, who

50. *See* Dianne Bartz, *Senators Criticize FTC's Reported Facebook Settlement*, REUTERS (Jul. 16, 2019), https://www.reuters.com/article/us-usa-facebook-ftc/senators-criticize-ftcs-reported-facebook-settlement-idUSKCN1UB25O.

51. Hannah Bloch-Wehba, *The Promise and Perils of Tech Whistleblowing*, 118 NW. L. REV. (forthcoming 2024) (manuscript at 11) (quoting Rory Van Loo, *The Missing Regulatory State: Monitoring Businesses in an Age of Surveillance*, 72 VAND. L. REV. 1563, 1565 (2019)).

52. *Id.*

53. *Id.* at 13–14.

54. *Id.* at 11–12.

55. *See, e.g.*, FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION 165 (2015) (maintaining that it was uncertain whether the FTC had the needed tools to effectively investigate allegations of "search bias").

blew a whistle on Twitter through the SEC in July 2022.[56] Twitter had already faced FTC action and, as a result, was obligated to comply with consumer data practice audits and interviews.[57] Zatko testified that the FTC's evaluations and examinations were essentially interview questions—accepting the companies' answers as truth—and were without quantifiable measurements.[58] Zatko's disclosures led to an FTC lawsuit, resulting in Twitter being ordered to pay $150 million in civil penalties.[59]

The FTC's interview practices, combined with the example above, suggest the FTC is also lacking in the technical expertise or resources to really understand the technology it regulates and certain problems that could or actually do result. Just five years ago, an exchange on the Senate floor demonstrated the leaps and bounds required to get the lagging technical expertise of regulators up to speed:

"[ORRIN] HATCH [Senator]: . . . [H]ow do you sustain a business model in which users do not pay for your service?

[MARK] ZUCKERBERG [CEO of Facebook]: Senator, we run ads.

HATCH: I see. That's great."[60]

Since this exchange, much more has been revealed and much has been learned about consumer data privacy practices. However, as revealed by Zatko in the summer of 2022, the technical expertise held by regulators is demonstrably lagging compared to highly

56. Mariam Baksh, *Whistleblower Explains How Twitter Easily Skirted FTC's Data Security Enforcement*, NEXTGOV/FCW (Sept. 14, 2022), https://www.nextgov.com/cybersecurity/2022/09/whistleblower-explains-how-twitter-easily-skirted-ftcs-data-security-enforcement/377130. The inadequacy of the SEC's whistleblower program for unfair or deceptive consumer data practices is discussed in Part IV. Generally, an insider is not eligible for a reward unless the SEC brings action. *See* 7 U.S.C. § 26(b)–(h) (setting forth whistleblower reward and protection requirements); 15 U.S.C. § 78u-6(b)(1) (discussing reward eligibility); 15 U.S.C. § 78u-6(a)(5) (including reward eligibility when there is a subsequent successful action "brought by an entity described in subclauses (I) through (IV) of subsection (h)(2)(D)(i)"); 15 U.S.C. § 78u-6(h)(2)(D)(i) (including "an appropriate regulatory authority"); SEC Rule 240.21F-4(g) (defining an "[a]ppropriate regulatory authority" as "an appropriate regulatory agency other than the Commission"); 15 U.S.C. § 78c(a)(34) (not mentioning the FTC).

57. Twitter, Inc., No. C-4316 (F.T.C. Mar. 2, 2011).

58. *Transcript*, *supra* note 9.

59. United States v. Twitter, Inc., No. 3:22-cv-3070, 2022 WL 1768852, at *1 (N.D. Cal. May 26, 2022).

60. *Transcript of Mark Zuckerberg's Senate Hearing*, WASH. POST (Apr. 10, 2018, 8:25 PM), https://perma.cc/7PFP-V3U5.

sophisticated commercial entities. More hurdles remain to bring a successful action with deterring effects.

C. The FTC's Enforcement Power Legal Process Is Also Procedurally Inefficient

Assuming an investigation is successful enough, the FTC then develops a record and takes legal action.[61] The FTC generally cannot seek civil (monetary) penalties in first actions against a company.[62] Rather, the FTC's first action against a particular company results in a court ruling imposing only equitable monetary remedies, or non-monetary imposed obligations.[63] Another option allows a company to preempt a court order and voluntarily enter into an agreement imposing obligations.[64] As mentioned above, typical FTC enforcement remedies require prohibiting acts complained of, remediating problematic acts, deleting wrongfully obtained information, modifying privacy policy, establishing a comprehensive privacy program, undergoing biennial audits, and adhering to record-keeping and reporting obligations.[65]

Even after an agreement becomes binding, thus leaving the company susceptible to civil penalties if found to be in violation of its agreement, there are still issues the FTC faces. The recent Twitter whistleblower case in the summer of 2022 illustrates one such problem.[66] The initial obligations imposed by the FTC after its first action against a company seek to foster a level of transparency and promote accountability through things like reporting requirements and audits.[67]

However, when Zatko testified in front of Congress about his experience working as Twitter's CSO, he revealed an unsettling reality: bound companies' reporting requirements are easily manipulated, leaving them with the ability to still practice unfair or deceptive practices without huge deterrence.[68]

61. *See* Solove & Hartzog, *supra* note 39, at 69.
62. *See supra* notes 38 and 39.
63. *See id.*
64. *See id.* at 610.
65. *See id.* at 614–18.
66. *See* Baksh, *supra* note 56.
67. *See* Solove & Hartzog, *supra* note 39, at 614–18.
68. *See* Baksh, *supra* note 56.

Zatko testified that Twitter did not admit to its systemic failure to delete consumer data when asked directly whether or not it did so in an interview by the FTC.[69] Instead, Twitter may have intentionally tried to mislead the FTC when it skirted the question and replied, "we deactivate users."[70] In reality, Twitter had not deleted user information.[71] Additionally, Zatko testified "third-party" FTC-required security certifiers were often hired by Twitter itself, indicating "a conflict of interest."[72] It was only through a whistleblower that the FTC eventually became aware of what was going on.[73]

The above discussion indicates the hurdles for the FTC are so numerous that their efficiency and deterrence is significantly hampered. A first action is effectively a warning, and though it has the potential for eventual huge consequences, companies recognize and take advantage of the uphill battle the FTC faces. As a result, companies have demonstrably found workarounds, see the risk-to-reward ratio as still heavily in their favor, and continue at least some of their unfair or deceptive practices with consumer data.

D. The FTC's Enforcement Power Can Be Inadequate, Resulting in Undeterred Companies with Compliers Facing Competitive Disadvantages

Aside from being inefficient, the FTC's enforcement power can also be inadequate. This could happen where actual harm has already occurred, and the FTC is only on its first action of enforcing with only non-punitive and equitable remedies available.[74] Companies are not deterred from ensuring consumer data is used and stored properly because of the FTC's level of inefficiency and inadequate remedies.[75] Without deterrence and with a lucrative alternative, prevalent risks to consumers' data follows.[76]

69. *Id.*

70. *Id.*

71. *Id.*

72. *Id.*

73. *Id.*

74. *See supra* note 38.

75. *See The U.S. Urgently Needs a Data Protection Agency*, ELEC. PRIV. INFO. CTR., https://epic.org/campaigns/dpa (last visited Oct. 6, 2023) (discussing the failures of the FTC to enforce privacy laws).

76. *See id.*

Inconsistent investment in compliance can also create inequitable competitive disadvantages among companies.[77] This happens when businesses trying to comply invest significantly in compliance.[78] As a result, their data practices may become less lucrative.[79]

A competitive disadvantage can be further exacerbated because of the lack of clarity in what warrants unfair or deceptive trade practices in the data privacy context.[80] The FTC has yet to use its rulemaking power to articulate bright line requirements or exclusions.[81] Although an abundance of rulemaking can be overburdening,[82] a lean and efficient level of bright line rules would help point companies in the right direction for issues the companies must address.[83] Although companies can look to prior actions by the FTC in this context, some situations are uniquely fact-specific, and it is more difficult to discern how narrowly or broadly the precedent may apply.[84] A desirable example that would not impede innovation would include a bright line to make clear where there would be carveouts or exceptions. Without bright lines or articulated exceptions, the disadvantage to companies striving to comply increases relative to those that continue to capitalize on non-compliance or to those that put forth minimal efforts.[85]

77. This Note focuses on the federal enforcement by the FTC and does not discuss the regulation of consumer data privacy at the state level. However, inconsistent investment in compliance and overburdening companies due to conflicts among patchwork state laws is a common concern at the state level. *E.g.*, DANIEL CASTRO, LUKE DASCOLI & GILLIAN DIEBOLD, THE LOOMING COST OF A PATCHWORK OF STATE PRIVACY LAWS 12–16 (2022) https://itif.org/publications/2022/01/24/looming-cost-patchwork-state-privacy-laws.

78. Trade Regulation Rule, *supra* note 16, at 51280.

79. *See id.*

80. Companies can be without clear notice because "[w]hat constitutes a deceptive or unfair trade practice has evolved, depending on what business practices the FTC has deemed problematic at any given time." Clark D. Asay, *Consumer Information Privacy and the Problem(s) of Third-Party Disclosures*, 11 NW. J. TECH. & INTELL. PROP. 321, 327 (2013).

81. *See* Trade Regulation Rule, *supra* note 16, at 51289.

82. A discussion of ex ante and ex post redress is discussed in Part IV. Specifically, an overly prescriptive approach with an abundance of bright line rules could be overburdening and hinder innovation because it would tend to assume a single set of consumer preferences or presume ways in which technology can or will evolve. The alternative, a broad ex ante approach, would scrutinize all possible data activities and violations in advance. It can also assume one general type of data preferences.

83. *See* Asay, *supra* note 80, at 327.

84. *See* Trade Regulation Rule, *supra* note 16, at 51280–81.

85. *See id.*

However, some might argue that the current level of inefficiency is acceptable and possibly desirable so that technological innovation may continue without being overly hindered.[86] This notion is rooted in the assumption that consumer data provides benefits to society at large that outweigh the level of current known harms resulting from inefficient data privacy regulation and enforcement.

A discussion of data privacy harms takes place below. It establishes that a lack of data privacy and security fostered by inefficient regulation enforcement has resulted in actual direct harms, grievous attenuated harm, and massive risks of additional kinds of known and lesser-known harm. The next Part reveals a hole and rebuts the above stated assumption that would prefer to promote the status quo: the argument for the status quo does not adequately give weight to the serious risk of not yet fully understood or seemingly attenuated harms. The following discussion furthers the notion that current inefficient enforcement is not acceptable, and that the proposed whistleblower incentive is needed. An incentive would promote efficiency and accountability, foster research and understanding, but refrain from over-burdening commercial interests.

III. KNOWN HARMS, LESSER-KNOWN HARMS, AND CHALLENGES IN ADEQUATELY ADDRESSING HARM

Addressing harm is difficult in the data privacy context. This is because being able to show legally cognizable harm is often a prerequisite for data privacy actions and an element for many causes of action.[87] This discussion broadly and briefly examines the range of harms from unfair or misleading consumer data practice and how those harms affect individuals, society, businesses, and national security. This is done to help explain important features of the proposed incentive.

First, there is a brief overview of known and concrete harms that result from a lack of data privacy. Second, there is a discussion about lesser understood harms that are difficult to cure through

86. Stephanie Comstock Ondrof, "*Senator, We Run Ads": Advocating for a US Self-Regulatory Response to the EU General Data Protection Regulation*, 28 GEO. MASON L. REV. 815, 848–52 (2021).

87. *See* Citron & Solove, *supra* note 12, at 796.

legal remedies but are important to understand and prevent. Third, this Part also discusses some of the good things happening in the data privacy context that can be further cultivated through the proposed incentive.

A. Known and Concrete Harms

"Harms involve injuries, setbacks, losses, or impairments to well-being[,]"[88] leaving people or society worse off. Regarding legally known and recognizable harm, data privacy breaches present tangible financial or physical harms ranging from mild to extreme.[89] Various forms of stalking from data breaches, improper collection, or data sales have led to multiple types of damages. These include death threats,[90] harassment,[91] "doxing,"[92] "swatting,"[93] and murder.[94] Data breaches from a lack of security or improper practices can also leave consumers more vulnerable to cyber-security threats, identity fraud, and theft, with financial repercussions ranging from slight to significant.

In November 2022, the FTC hosted its seventh annual PrivacyCon, where the harm of algorithmic bias was a key focus.[95] A panel on *Bias in Algorithms* presented researchers' findings:

88. *Id*. at 799.

89. *See id*. at 830–61.

90. *See Planned Parenthood of the Columbia/Willamette, Inc. v. Am. Coal. of Life Activists*, 290 F.3d 1058, 1065 (9th Cir. 2002) (en banc) (involving abortion doctors and their families receiving death threats).

91. See Citron & Solove, *supra* note 12, at 818; *see generally* DANIELLE KEATS CITRON, HATE CRIMES IN CYBERSPACE (2014) (providing an in-depth examination of the prevalence and extent of personal cyber-attacks, including harassment, and proposing practical, lawful ways to prevent and punish online harassment).

92. "Doxing" is a form of intimidation where data is revealed to help others in tracking a person down to attack, threaten, or otherwise harass them. *See* Ryan Goodrich, *What Is Doxing?*, TECH NEWS DAILY (Apr. 2, 2013) https://web.archive.org/web/20141029095609/http://www.technewsdaily.com/17590-what-is-doxing.html.

93. "Swatting" is a harassment practice that involves falsely calling in an emergency threat to law enforcement to send officials to an address. *See* Daniel J. Solove, *Data is What Data Does: Regulating Based on Harm and Risk Instead of Sensitive Data*, 118 NW. U. L. REV. (forthcoming 2024) (manuscript at 36) [hereinafter Solove, *Data is What Data Does*].

94. An example of this is described in *Remsburg v. Docusearch*, 816 A.2d 1001 (N.H. 2003), a case which arose after a stalker bought a woman's work address from a personal data search company, then stalked and murdered her.

95. *See* Ziad Obermeyer, Brian Powers, Christine Vogeli & Sendhil Mullainathan, *Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations*, 366 AM. ASS'N FOR ADVANCEMENT SCI. 447, 447 (Oct. 25, 2022).

an algorithm used with the intentions of targeting medical interventions to the sickest patients ended up directing resources to a healthier, white population.[96] This same group of researchers established that black patients were sicker than white patients, with black patients needing an increase in help from 17.7–46.5% to alleviate the disparity.[97] This research is unsettling and indicates the prevalence of harms arising from the data collection and algorithms used in everyday life.

B. Issues with Less-Cognizable Privacy Harms

Although recognizable harms are potentially pervasive, data privacy harms more often involve injuries that are difficult to assign a monetary value to and are thus less legally recognizable. These injuries range from mild to monumental in either isolation or aggregation. For example, it would be difficult to quantify a monetary remedy for the mental anguish teen girls experience through Instagram flooding their feed with targeted squares displaying body image perfectionism. Even though an insider leaked research that showed Instagram made one in three teen girls feel worse about their bodies, the current legal scheme would

96. Algorithms and privacy are often treated as separate and distinct areas for regulation. Algorithms can include "design harms," which are considered separate and distinct from "privacy harms." To illustrate the difference, an Instagram or other camera application filter that uses teens' faces—making their faces appear to have bigger lips, eyes, etc.—can promulgate body dysmorphia. This could be considered a design harm. These filters can also be gamified to encourage sharing the filter with friends, expanding the impact. Alternatively, a privacy harm could exist in this same scenario where that same filter captures and stores the facial features of teens, tracks who the filter is shared with, and subsequently uses that data to further develop harmful products or promulgate ads with similar issues to vulnerable individuals. This example demonstrates that a privacy harm can exacerbate a design harm. Algorithms and artificial intelligence continue to develop by the data received and learned from. As such, algorithms present situations in which the data used could result in an area of overlap between a design and a privacy harm. An algorithm's use of consumer data could amount to a deceptive or unfair use of the collected data. This Note does not focus on design harms, however, for further discussion of the two and their overlap, see *Harmful Design in Digital Markets: How Online Choice Architecture Practices Can Undermine Consumer Choice and Control over Personal Information*, DIGITAL REGULATION COOPERATION FORUM, https://www.drcf.org.uk/__data/assets/pdf_file/0024/266226/Harmful-Design-in-Digital-Markets-ICO-CMA-joint-position-paper.pdf (last visited Oct. 10, 2023).

97. Obermeyer et al., *supra* note 95, at 447.

struggle to quantify this harm on an individual basis.[98] On an aggregated scale, the harm is huge and the societal impact through its ripple effect is strong. But articulating legal restoration in cases like these is a practical struggle.

These harms are difficult to quantify but nonetheless carry consequences considered repugnant in other areas of the law. Broken contractual promises occur when consumers consent to company data privacy practices and companies stray beyond these condoned data uses. This also results in thwarted expectations about how people's data will be collected, used, disclosed, and traded. The downstream effects of the broken promises, thwarted consumer expectations, and unsanctioned data use have harmful consequences that are difficult to determine and often impossible to calculate.[99] The biased algorithms are just beginning to be understood and so are still somewhat incognizable at law. Harms range from unwanted spam or an eerie sense of surveillance to widespread discrimination, exacerbation of racial inequities, political manipulation, shame, embarrassment, ridicule, reputational humiliation, emotional distress, anxiety, and depression.[100]

Articulating "harm" in this context is difficult because some harms are developing, just like the technology the harms stem from.[101] One other fundamental reason may be best illustrated using an analogy to a once omnipresent and largely unregulated industry—the tobacco industry.

"Facebook seems to be taking a page from the textbook of Big Tobacco—targeting teens with potentially dangerous products while masking the science in public."[102] Tobacco was a substance that was initially unregulated, and its harms were largely not understood.[103] It was highly popular, highly lucrative, and even touted to be beneficial.[104] However, it was eventually discovered

98. *See* Georgia Wells, Jeff Horwitz & Deepa Seetharaman, *Facebook Knows Instagram Is Toxic for Teen Girls, Company Documents Show*, WALL ST. J. (Sept. 14, 2021, 7:59 AM), https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739; Citron & Solove, *supra* note 12, at 796–813.

99. *See* Citron & Solove, *supra* note 12, at 797.

100. *See id.*

101. *See id.* at 817–18.

102. Wells et al., *supra* note 98.

103. *Id.*

104. *See* Anne Charleton, *Medicinal Uses of Tobacco in History*, 97 J. ROYAL SOC'Y MED. 292, 292 (2004).

that tobacco has serious, proximate, and concrete harms.[105] Before thorough research and understanding, a harm like heart disease to smokers or eventual birth defects resulting from secondhand smoke exposure may have at first seemed too attenuated. But with more study, it eventually became clear these harms were directly linked to the consumption of tobacco.[106]

The important difference between data and drugs illustrates one fundamental problem in adequately addressing harm: Consumer data is unlike tobacco, because scientists could simply purchase and deconstruct a cigarette, then proceed to test its contents and effects in various ways. With data, private companies keep their data, inferences, algorithms, and hyper-specific uses under lock and key.[107] Because of this, harms are much more difficult to research, understand, and discover. As a result, many data privacy harms will continue to be without effective protections and remedies through the law.

Companies have developed restrictive policies on voluntarily sharing information.[108] Even if researchers went ahead and purchased consumer data from entities, researchers would still be unable to fully understand how it is used or what kind of correlations are possible without insight.[109] Researchers can only softly predict what correlations exist between the data, the amount of harm, and the types of harm that result.[110] Without insight, researchers and regulators lack the ability to develop a first-hand understanding or otherwise deconstruct the gathered data, its use, and its harms.[111]

Pressing questions to alleviate known issues cannot be answered without insight into companies' proprietary data. For example, are there patterns in how news outlets and social media companies filter, censor, or promote the news to consumers by ideology, race, or other discriminatory factors? How does consumer

105. *See id.*

106. Press Release, Center for Disease Control, Smoking Early in Pregnancy Raises Risks of Heart Defects in Newborns (Apr. 7, 2008).

107. *See* Mary D. Fan, *The Right to Benefit From Big Data as a Public Resource*, 96 N.Y.U. L. REV. 1438, 1438–92 (Nov. 2021) (discussing the value and benefit the public would have with a right to access collected data as a public resource).

108. *See id.* at 1443–44.

109. *See id.*

110. *See id.*

111. *See id.*

data influence or relate to the rise of violent extremism? How might consumer data be used to discriminate access to health care, employment, price-savings, or other opportunities?

The law's shortcomings in remediating, preventing, and even understanding harms are exacerbated because companies are without motivation to research concrete and less-cognizable harms. However, there are important instances where companies voluntarily gave insight to help solve societal issues.

C. Pieces of Positives to Build on in Data Privacy Law

There are instances of important, good things going on in the data privacy context that should continue to be incentivized and built on. For example, in 2020, Google voluntarily released free data sets to help teach machine learning how to help prevent the spread of COVID-19.[112] Additionally, Facebook and Social Science One released social science datasets to study elections and the spread of misinformation.[113] These are instances that should be incentivized and facilitated through the law. As for the regulatory system, the FTC has proven to be an important tool to help prevent potential data privacy damages. Incentives should foster companies coming forward with research that provides valuable benefits to society and diminishes consumer harms.

The FTC has a particular advantage in regulating consumer data because it does not have the requirement to establish harm for standing or damages.[114] For example, in October 2019, the FTC brought its first case against developers of "stalking" apps.[115] The apps were designed to run secretly in the background—with the ability to hide the app icon so a phone user would not know it was installed.[116] These apps were uniquely suited to illegal and

112. *See* Kyle Wiggers, *Google Cloud Releases Covid-19 Data Sets to Foster Coronavirus-Fighting AI Models*, VENTUREBEAT (Mar. 30, 2020, 9:39 AM), *https://venturebeat.com/2020/03/30/google-launches-covid-19-public-datasets-program-to-foster-coronavirus-fighting-ai-models*.

113. Gary King & Nathaniel Persily, *Unprecedented Facebook URLs Dataset Now Available for Academic Research Through Social Science One*, SOC. SCI. ONE (Feb. 13, 2020), https://socialscience.one/blog/unprecedented-facebook-urls-dataset-now-available-research-through-social-science-one.

114. *See* Citron & Solove, *supra* note 12, at 814.

115. *See* Retina A-X Studios, LLC v. James N. Johns, Jr., No. C-4711 (F.T.C. Oct. 2019).

116. *See id.*

dangerous uses.[117] The apps allowed purchasers to monitor the mobile devices on which the apps were installed, without the knowledge or permission of the device's user.[118] The case resulted in a settlement, aimed to resolve allegations that these apps "compromised the privacy and security of consumer devices on which they were installed."[119] Without the FTC's action, an ordinary consumer would have a difficult time proving monetary damages based on the app's information gathering alone without proof of monetary or other legally recognizable harm.

In sum, even though there are instances of progress, the argument for the status quo does not adequately give weight to the serious risk of not yet fully understood or seemingly attenuated harms. Nor does it account for the inability to better understand them. The current inefficient enforcement is not acceptable, and the proposed whistleblower incentive is needed. This is because it promotes efficiency and accountability, fosters research and understanding, and can prevent harm, but refrains from over-burdening or over-regulating commercial interests.

The next Part shows how a whistleblower incentive would offer unique solutions to an area of law that struggles to both recognize and remedy harm. Needed incentive traits are discussed below.

IV. CARROT CHARACTERISTICS

Incentives are an integral part of intellectual property law. For example, patent law works to discourage trade secrets by providing inventors with an incentive to publicly disclose their original and substantial inventions to the United States Patent and Trademark Office.[120] In exchange, if a patent is granted, inventors can secure exclusive rights to their discoveries for a limited time.[121] An incentive in data privacy law could facilitate commercial progress and discourage total trade secrecy by rewarding beneficial research

117. *See id.*

118. *See id.*

119. *USA: FTC Announces Proposed Settlement with Retina-X Studios for Tracking Apps*, DATAGUIDANCE (Oct. 23, 2019), https://www.dataguidance.com/news/usa-ftc-announces-proposed-settlement-retina-x-studios.

120. *See* Anderson, *supra* note 15, at 921.

121. *See* U.S. CONST. art. I, § 8, cl. 8; *see also* Wendy J. Gordon, *Intellectual Property*, *in* THE OXFORD HANDBOOK OF LEGAL STUDIES 617, 632 (Peter Cane & Mark Tushnet eds., 2003).

or access to data. An incentive would also remain rooted in the notion of protecting consumers.

This Part outlines and analyzes four necessary qualities of a whistleblower incentive relating to the FTC and consumer data privacy: it should (1) be narrow and only address unfair or misleading consumer data practices; (2) provide protection for whistleblowers with an avenue for remaining anonymous and prohibiting companies from retaliating against protected whistleblowers; (3) include clear-cut requirements for a whistleblower to be eligible for both protection and reward; and (4) include a mandatory reward when requirements are satisfied, with a specific range the reward amount will be based on. The supporting reasons for each of the above-listed characteristics are discussed below.

A. A Narrowly Tailored Incentive

First, a whistleblower incentive should be narrow and address only consumer data. Insiders would report to the FTC either (1) known deceptive or unfair data practices that would violate the law or a previously entered agreement with the FTC; or (2) practices that would likely be considered "deceptive" or unfair. The incentive should be tailored to this purpose for multiple reasons. Having a clear-cut and narrow incentive will provide certainty for tech workers as they go through their own complicated cost, risk, and ethical analysis.[122] This clarity helps tech workers know whether they would be protectable whistleblowers or potentially illegal leakers facing hefty financial liability or even criminal charges. Clarity and certainty also help fight other forms of deterrence from disclosure that tech workers must grapple with. Specifically, a potential whistleblower has real relationships with

122. There is a current bill in the House of Representatives, H.R. 6093, which proposes a broad whistleblower incentive for the FTC generally, but it has multiple flaws that will likely keep it from passing. Each of the characteristics discussed in this Note are largely not present in H.R. 6093. Introduced in November 2021, this bill is set to expire if not voted on within two years, by November 2023. Though not discussed further, the failure of this bill helps demonstrate that an incentive needs to be narrower. *See* H.R. 6093, 117th Cong. (2021) See *also* Dallas Hammer & Jason Zuckerman, *FTC Whistleblower Act Would Reward and Protect Whistleblowing About Data Privacy Misconduct and Other Deceptive Practices*, THE NATIONAL LAW REVIEW (Dec. 3, 2021), https://www.natlawreview.com/article/ftc-whistleblower-act-would-reward-and-protect-whistleblowing-about-data-privacy (analyzing the characteristics of H.R. 6093).

friends and colleagues at stake. A worker's reputation and career are also on the line because "[i]t's hard to find employment when you've been branded as the whistleblower."[123]

Additionally, the clear parameters will make things more manageable for the FTC by providing better quality insight. Funneling a whistleblower incentive to only data privacy will enable the FTC to develop a more focused, meaningful, and calculated response to the insight it receives. One important reason for a whistleblower incentive is to offer otherwise unavailable insight for policymakers to learn from. This focused insight will help the FTC as it develops and promulgates new rules. Insight and learning are needed to prevent an overly cumbersome number of rules from being implemented. Instead, learning more from companies will help regulators foster lean and efficient guidelines for what constitutes unfair or deceptive data practices. With the insight and resulting rules, the FTC will be able to improve its efficiency in significant ways.

Narrow parameters are also necessary to balance and pace the effects on the data industry. A moderate pace is ideal because an essential part of the FTC's purpose is to avoid unduly burdening legitimate business activity. Historically, the FTC has taken a stair-stepping and slow approach to developing its own sort of precedent.[124] This is done intentionally to give companies more notice of what is expected from them without requiring the FTC to go through a formal rulemaking process.[125] Further, a tailored incentive will also help companies conceal their otherwise legal trade secrets. This will help insiders limit what they disclose, which also helps promote commercial interests.

B. An Incentive with an Avenue for Anonymity and Protection

Second, a whistleblower incentive should provide protection by offering an avenue for anonymity to insiders. If anonymity is not chosen, there should be provisions prohibiting companies from retaliating against protected whistleblowers. Anonymity and

123. Bloch-Wehba, *supra* note 51 (manuscript at 30) (citing Kristian Hernandez, *COVID Underscores Lack of Whistleblower Protections*, STATELINE (Feb. 14, 2022), https://pew.org/3oHpHaR).

124. *See* Trade Regulation Rule, *supra* note 16, at 51273.

125. *See id.*

protection from retaliation are necessary to help alleviate some of the social costs involved that are outside of the civil, criminal, and professional liability. A real-life situation illustrates the need for such provisions below.

A leak by an Apple employee illustrates the stakes insiders face when forced with difficult ethical questions. An Apple employee leaked a memo posted on the company's internal blog.[126] Apple's memo offered statistics on employee leaks and enforcement.[127] It reminded workers that "[i]n 2017, Apple caught 29 leakers. 12 of those were arrested."[128] The memo also touted that the company was able to pinpoint and catch a single employee who leaked information from a meeting with hundreds of employees in attendance.[129]

Apple was not just attempting to ensure total trade secrecy; it was intimidating employees by threatening that it would take criminal action against employees who leaked insider information. In addition to criminal penalties, Apple's employees would also face legal and social consequences. Insider insight is needed in an information imbalanced system, but the stakes are high.

Protection from professional retaliation is also needed because the foundation of at-will employment allows employers to fire their employees "for what [employees] say" or for most other reasons.[130] Additionally, the constitutional freedom of expression is not enough to protect private-sector employees from their "voluntary" non-disclosure agreements (NDAs) or other acceptable employment practices.[131]

Further, without retaliation protections, whistleblowers could face both civil and criminal liability. Freedom of contract often prevails over freedom of employee speech. Contracts are used to

126. *See* Mark Gurman, *Apple Warns Employees to Stop Leaking Information to Media*, BLOOMBERG (Apr. 13, 2018, 11:18 AM), https://www.bloomberg.com/news/articles/2018-04-13/apple-warns-employees-to-stop-leaking-information-to-media.

127. *See id.*

128. *Id.*

129. *Id*.

130. *See* Bloch-Wehba, *supra* note 51 (manuscript at 15) (quoting Charlotte Garden, *Was It Something I Said?: Legal Protections for Employee Speech* 1, ECON. POL'Y INSTITUTE (2022)).

131. Bloch-Wehba, *supra* note 51 (manuscript at 16); *see also* Pauline T. Kim, *Market Norms and Constitutional Values in the Government Workplace*, 94 N.C. L. REV. 601, 610 (2016).

restrict trade secrets. An employee divulging trade secrets can face criminal charges in some instances.[132]

Employee NDAs can also extend beyond legitimate trade secrets to broadly defined "confidential information."[133] Even when employers know their NDAs are so overbroad that they are legally unenforceable, companies use them anyway. This is because employees are still deterred from going anywhere close to a line that would cause legal complications from an NDA.[134]

Whistleblowers also need protection because even if they want and try to remain anonymous, there is a legitimate chance whistleblowers will get caught by their company. As the example of Apple's internal memo illustrates, companies are good at catching employee leaks due to a broad use of workplace surveillance. Companies prioritize intense investigative efforts to catch and punish leakers. Without some level of protection, whether there is a reward on the line, the numerous risks and costs employees would otherwise face will almost always outweigh their desire and ability to inform the public, academics, and regulators of wrongdoing.

C. An Incentive with Clear-Cut Requirements for Reward and Protection Eligibility

Third, a whistleblower incentive should include clear-cut requirements and exclusions for an employee to be eligible for both protection and reward. Some ideal requirements and exclusions would include:

- An FTC Whistleblower must voluntarily provide original information relating to a possible violation of a business conducting unfair or deceptive consumer data practices.
- This insight must lead to the successful enforcement by the FTC of a federal court or administrative action in which the FTC obtains monetary sanctions totaling more than one million dollars.

132. *See* 18 U.S.C. §§ 1831–39; *see also* Madeleine Cane, Michael Bednarcyzk, Maxwell Nides, Patrick Engle & Quinlan Cummings, *Intellectual Property Crimes*, 58 AM. CRIM. L. REV. 1151, 1154–55 (2021).

133. Orly Lobel, *NDAs Are Out of Control. Here's What Needs to Change*, HARV. BUS. REV. (Jan. 30, 2018), https://hbr.org/2018/01/ndas-are-out-of-control-heres-what-needs-to-change.

134. *See id.*

- Whistleblowers would still be eligible for a reward if they first report internally and the company informs the FTC about the violations.
- Certain people would not generally be eligible for rewards, including people who have pre-existing legal or contractual duties to report their information to the FTC, such as attorneys, compliance personnel, and internal auditors.
- To fall under the FTC's employer anti-retaliation protection, the whistleblower must possess a reasonable belief that the information that the whistleblower provides relates to a possible FTC violation that has occurred, is ongoing, or is about to occur.

Different schools of thought provide legitimate reasons for the clear-cut protection and reward eligibility requirements. As established in the discussion of the FTC's shortcomings, there is a legitimate need to help with enforcement because the FTC is without sufficient human-capital, monetary-capital, and lacks technological detection ability.[135]

There is also a conceptual regulatory need. The FTC's current actions do not create bright-line precedent to follow:[136] Instead, clearer regulations should be formed to synthesize and streamline the holdings of FTC caselaw. Whistleblowers will help point regulators in the right direction of problems. This gives regulators insight on common problems, how to create clearer rules, and where exceptions need to be expressly carved out.

Clear-cut protection and incentive requirements are also desirable because such requirements are already proven to work in other areas of law—namely, securities violations. However, the SEC's whistleblower laws do not fully cover unfair or deceptive data practices. Data practices are largely outside the SEC's enforcement power and are instead within the FTC's jurisdiction. The specifics are discussed below.

The SEC's Dodd Frank Act helps show that incentives are a proven method.[137] Since 2010, whistleblower tips have helped the SEC to recover nearly $5 billion in monetary penalties, with $1.3

135. *See* Bloch-Wehba, *supra* note 51 (manuscript at 50–51).

136. *See* Solove & Hartzog, *supra* note 39, at 585–86.

137. *See* 7 U.S.C. § 26(b)–(h) (setting forth whistleblower reward and protection requirements).

billion in investor restitution.[138] These instances would likely be undiscoverable without a whistleblower pointing a finger in the right direction. The Commodities Futures Trading Commission (CFTC) is an additional example of an agency with a successful whistleblower program.[139] The CFTC issued its first reward from its Whistleblower Program in 2014 and has since granted several whistleblower awards, which have led to cumulative monetary sanctions of more than $3 billion.[140]

The SEC's whistleblower program also establishes that there are other reasons for an FTC whistleblower program regarding unfair or deceptive data privacy practices. This is because there are already important instances where tech workers, in the absence of specific data privacy whistleblower protections, have innovatively situated themselves to fall within the SEC's whistleblower protection laws.[141] In some of these instances, the FTC used information gained through the whistleblowers' leaks to the SEC to subsequently bring successful FTC actions against the companies. However, although helpful for the FTC and for insiders seeking protection, the SEC's whistleblower laws are not an obvious route to protection for insiders when it comes to data security. Without a clear path and relatively certain protections, the vast majority of appropriate insider insight will remain chilled. A narrow FTC whistleblower incentive and protection, on the other hand, would make things clearer and facilitate appropriate insight.

It also seems that an SEC whistleblower would not be eligible to recover any monetary bounty where only the FTC brings subsequent successful action. This is because the FTC is not listed under the SEC's laws that articulate the requirements for a whistleblower to be eligible for a reward.[142] There is no mention of

138. SEC. & EXCH. COMM'N, SEC ANNUAL REPORT TO CONGRESS: WHISTLEBLOWER PROGRAM 1 (2021).

139. *See id.*

140. Press Release, Commodities Futures Trading Commission, CFTC Awards Approximately $10 Million to a Whistleblower (Mar. 18, 2022).

141. *See supra* Section IV.A.

142. *See* 15 U.S.C. § 78u-6(b)(1) (discussing reward eligibility); 15 U.S.C. § 78u-6(a)(5) (including reward eligibility when there is a subsequent successful action "brought by an entity described in subclauses (I) through (IV) of subsection (h)(2)(D)(i)"); 15 U.S.C. § 78u-6(h)(2)(D)(i) (includes "an appropriate regulatory authority"); SEC Rule 240.21F-4(g) (defining an "[a]ppropriate regulatory authority" as "an appropriate regulatory agency other than the Commission"); 15 U.S.C. § 78c(a)(34) (not mentioning the FTC).

the FTC or use of sweeping terminology that would otherwise qualify the FTC in the SEC's rules defining "appropriate regulatory agency" or "appropriate regulatory authority."[143] Thus, an FTC whistleblower incentive would again provide a more certain and clear path for insiders.

D. An Incentive with a Mandatory and Specific Reward Range

Fourth, a whistleblower incentive should, when requirements are satisfied, include a mandatory (as opposed to discretionary) reward with a specific range that the reward amount is based on. This point pairs with the benefits of clear lines and ease of risk analysis discussed in Part III; however, there is an additional important reason for mandatory rewards. Imposing mandatory rewards where requirements are met means the FTC would recover and benefit from monetary penalties. With the incentive in place and assuming a higher volume of successful enforcement actions, the FTC's portion of the recovery could be used to help fund its inadequate resources.

With these characteristics in mind, the next Part delves into eight benefits a narrow FTC whistleblower incentive would provide, including important ways it would improve deficiencies and inadequacies in the current FTC enforcement and regulatory system.

V. EIGHT ADVANTAGEOUS BITS OF THE CARROT BYTE

Tech leaks have already established their value and their necessity. Leaks have unearthed some of the most significant technology policy issues in the current landscape. Without these leaks, some important issues could have remained only suspicions and virtually undiscoverable. The uncovered issues' ripple effects have prompted federal regulatory action by the FTC, spurred congressional committees to convene,[144] and helped facilitate state efforts to get involved.[145] A few examples are included below.

143. *See id.*

144. *See* Bloch-Wehba, *supra* note 51 (manuscript at 23).

145. *See* Allison Slater Tate, *Facebook whistleblower Frances Haugen says parents make 1 big mistake with social media*, TODAY (Feb. 7, 2022, 5:06 PM), https://www.today.com/parents/teens/facebook-whistleblower-frances-haugen-rcna15256; Press Release, Utah Gov., Utah Sues TikTok Over Child Addiction Harm, Targets "Enmeshment" With Its China-Based Parent Company (Oct. 10, 2023), https://governor.utah.gov/2023/10/10/utah-sues-tiktok-over-child-addiction-harm-targets-enmeshment-with-its-china-based-parent-company.

In one notable example, Christopher Wylie, a former Cambridge Analytica employee, revealed how his company was using a new version of political campaigning through psychological targeting. His insight exposed that hundreds of thousands of users were lured to take a personality test—being paid to do so—and agreed to having their data collected and analyzed for "academic use."[146] The app "thisisyourdigitallife" was the test's platform. The app harvested data from both its users and their friends. At that time, Cambridge Analytica was run by Trump strategist Steve Bannon, and the company "used the harvested data to target [app] users with political advertising."[147] Although there was already suspicion surrounding Cambridge Analytica and Facebook's inappropriate data use and political targeting, both companies had been actively undermining such claims. The companies instead claimed nothing inappropriate had occurred and did not admit to exploiting collected user data without users' knowledge on behalf of political candidates. Without Wylie's leak that included supporting evidence, the suspicions could have dissipated and been disregarded.

Regarding election fraud and misinformation, Facebook leakers used company documents to establish that numerous employees repeatedly tried to raise red flags about misinformation spreading and conspiracies gaining traction before and after the contested November 2020 presidential election.[148]

In a separate instance, after Facebook allegedly dissolved its civic integrity team, a Facebook insider began taking the steps to intervene.[149] A Facebook insider facilitated a groundbreaking story by disclosing tens of thousands of internal Facebook documents to Congress and the Securities and Exchange Commission (SEC).[150] Another insider shared Facebook's research that revealed unsettling

146. Carole Cadwalladr & Emma Graham-Harrison, *Revealed: 50 Million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach*, THE GUARDIAN (Mar. 17, 2018, 6:03 PM), https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election.

147. Bloch-Wehba, *supra* note 51 (manuscript at 36).

148. Ryan Mac & Sheera Frenkel, *Internal Alarm, Public Shrugs: Facebook's Employees Dissect Its Election Role*, N.Y. TIMES (Oct. 25, 2021), https://www.nytimes.com/2021/10/22/technology/facebook-election-misinformation.html.

149. Billy Perrigo, *Inside Frances Haugen's Decision to Take on Facebook*, TIME MAG. (Nov. 22, 2021), https://time.com/6121931/frances-haugen-facebook-whistleblower-profile.

150. *Id*.

statistics demonstrating that "[t]he tendency to share only the best moments, a pressure to look perfect and an addictive product can send teens spiraling . . . ."[151]

Other tech giants have had similarly important leaks. After a Google executive side-stepped a critical question in his congressional testimony, Google leakers helped break the news that the company was developing a censored search engine for China.[152] With these notable examples in mind, this next Part delves into eight separate benefits a narrow FTC whistleblower incentive would provide.

A. This Incentive Would Provide a Balanced Approach to Policymaking

First, a carrot and its characteristics provide a palatable policy balance because it would not limit either the direction of industry innovation or the options available to consumers. The incentive provides a balanced policymaking approach because an incentive combined with the FTC as an enforcer would be a type of ex post redress. Ex post redress means unfair or deceptive data practices are reviewed and fines or other penalties are put into place after they have occurred.[153] An ex post redress refrains from becoming an "overly prescriptive approach that assumes only one set of [consumer] preferences . . . ." Nor would an incentive "presume[]. . . ways in which technology can [or will] evolve."[154] However, it would still facilitate authority intervention in a more efficient way when necessary to prevent harm.[155]

The alternative, a broad ex ante approach, would "scrutiniz[e] all . . . possible data activities and violations in advance." An ex ante approach can also assume one general type of data preferences. One example of a prescriptive ex ante approach is Europe's General Data Protection Regulation (GDPR), which has resulted in some products, like email management and online game

151. Wells et al., *supra* note 98.

152. Alexis C. Madrigal, *Silicon Valley Sieve: A Timeline of Tech-Industry Leaks*, THE ATLANTIC (Oct. 10, 2018), https://www.theatlantic.com/technology/archive/2018/10/timeline-tech-industry-leaks/572593.

153. *See* Jennifer Huddleston, *A Primer on Data Privacy Enforcement Options*, AM. ACTION F. (May 4, 2020), https://www.americanactionforum.org/insight/a-primer-on-data-privacy-enforcement-options/#ixzz7xD3qbjBT.

154. *Id.*

155. *Id*.

applications, becoming unavailable in Europe "due to the cost or even impossibility of compliance" with the GDPR.[156] Although the ex post approach alone may not be enough, it could help ensure any type of incorporated ex ante approach remains lean and efficient.

B. This Incentive More Directly Addresses Data Use, Harms, and Risk of Harms than the Current "Sensitive Data" Approach

Second, this incentive offers an enforcement tool to directly address use, harms, and risks of harms more efficiently than the FTC is currently able to. In Professor Daniel J. Solove's forthcoming article titled *Data is What Data Does: Regulating Based on Harm and Risk Instead of Sensitive Data*, Solove thoroughly addresses fundamental flaws with the current popular prescriptive approach that is framed around regulating "sensitive data."[157] In sum, he argues that regulations prohibiting broad and inconsistently defined "sensitive data" are unworkable for companies, counterproductive to innovation, and still pose serious risks of harms to consumers.[158]

The risks of harm remain due to notable omissions in sensitive data definitions, proxies that carry out harm without using "sensitive data," and because harmful algorithmic inferences based on "non-sensitive data" are highly accurate and proven to exist.[159] Instead, Solove argues for a system that would more directly address use, risk, and harm, and he notes the difficulty in this task.[160] He mentions a need for regulators to implement an approach that is proportionate to use, risk, and harm instead of a blanket approach with inconsistent definitions that have unclear boundaries.[161]

By being primarily geared toward an ex post redress, the incentive would give insight that could be evaluated based on the data's actual use. This would enable enforcement that is proportionate to the use of the data relating to harm or risk of harm. It would avoid creating a blanket approach that could otherwise

156. *Id.*

157. *See* Solove, *Data is What Data Does*, *supra* note 93.

158. *Id.*

159. *Id.* (manuscript at 33–43).

160. *Id.* (manuscript at 43–49).

161. *See id.* (manuscript at 29–30; 43–48).

treat uses disproportionately to their harm or risk of harm. This approach to addressing risk of harm also fosters a hybrid between an ex post facto and ex ante methodology. This is because regulators will gain valuable insight, learn from it, build on their expertise, and be in a better position to craft lean ex ante regulations. Equally as important, the incentive fosters enforcement by the FTC that would enable regulators to address harms no matter what type of data is used. Each of the later points of this Part address other inefficient aspects the FTC faces that were discussed in Part I.

Using the FTC to regulate harm in data privacy has another distinct advantage because actual harm is not a requirement for the FTC to have standing.[162] This removes the hurdle of having to show harm in a system that struggles with adequately addressing injuries in the data privacy context. The proposed whistleblower incentive would not change the way the FTC brings actions. As a result, a whistleblower incentive would not overreach into addressing harms that the system is not yet built to address, nor would it overburden commercial interests.

C. The Incentive Would Help Level the Current Data Information Imbalance, Inform Future Policymaking, and Easily Integrate into the Current System

Third, this incentive would be a counterweight to help balance the highly secretive nature of data information flow in a lagging regulatory system.[163] As discussed in section II(B), detectability is a major issue in determining when to bring initial actions, knowing where to look in investigations, understanding the evidence and technology being investigated, and keeping companies accountable to their subsequent settlement agreements. An incentive would facilitate insider insight that could provide valuable detection and evidence that is otherwise difficult to acquire. Also, where regulators are ill-equipped to fully understand what they are evaluating, whistleblowers can supply understanding and point regulators in the right direction.

162. *See* Asay, *supra* note 80, at 327 (citing PETER SWIRE & SOL BERMANN, INFORMATION PRIVACY 70 (2007)).

163. *See* Bloch-Wehba, *supra* note 51 (manuscript at 1).

Companies would also be less inclined to manipulate their answers to skirt liability throughout investigations.

Fourth, this incentive presents a palatable piece for the FTC to build on and learn from as the FTC confronts challenges in rulemaking. According to one report released by the agency, the FTC has brought more than 130 spam and spyware cases, 80 general privacy cases, 70 data security cases, and 100 cases involving Fair Credit Reporting Act violations.[164] Because of its extensive experience and congressional authority, the FTC is best situated to develop formal rules that create a balanced approach to consumer data privacy. A whistleblower incentive would provide an avenue for further insight, learning, and development. In the long run, by providing insight to the FTC and to Congress, regulations can be refined and adapted to better address risks and prevent injuries.

Fifth, a narrow incentive can be easily grafted onto any current or future "bundle of sticks" regulations. This advantage is straightforward and somewhat simple, but it should not be understated. Without cohesive federal laws that preempt patchwork state data privacy laws, adding any data privacy federal regulation quickly gets complicated even before considering politics. A whistleblower incentive would not add more regulations but rather would be used to assist current and future regulatory enforcement.

D. The Incentive Would Foster Transparency, Encourage Consistent Company Investment in Conformity, and Ultimately Promote Deterrence in an Efficient Way

Sixth, a whistleblower incentive could uniquely promote transparency and foster a better understanding of harm. Companies who might face liability as a result of a whistleblower could combat an extent of that liability by first making research of data use and its harms a priority; and second, by either making that research accessible to some level of outsiders or by allowing regulators, academics, and scientists to perform their own research with companies' gathered data. By doing this, the FTC and courts could take these acts into consideration when determining the extent of obligations imposed through settlement agreements or the amount sought in punitive recovery.

164. FED. TRADE COMM'N, PRIVACY AND DATA SECURITY UPDATE: 2019 (2019).

Seventh, the incentive encourages consistent investment by companies in implementing fair consumer data practices and following regulations. As discussed in Part II, the current lack of deterrence can create inequitable competitive disadvantages among companies and advantages for non-compliers. There are general principles that can be synthesized through the "precedent" created from previous FTC actions.[165] However, without clear rules and definitions, companies striving to comply can face hefty costs interpreting and complying with FTC precedent. With the threat of protected informants who are incentivized to come forward with a reward, companies will take their data privacy practices more seriously and make proper practices a priority. This incentive will bring equity where competitive disadvantages arise. Also, over time, the FTC will be able to bring more clarity to what warrants unfair or deceptive trade practices in the data privacy context through rulemaking and exclusions or consistent case law.

Eighth, a whistleblower incentive offers a way to promote deterrence without pumping more capital into enforcement agencies. As discussed in Part II, the FTC faces inadequate funding and minimal human capital to the point that it cannot facilitate a desirable amount of deterrence. This undermines its efficiency and ability to address harms. A whistleblower incentive would promote deterrence without more capital in multiple ways. Put simply, companies would be more deterred because the risk of FTC actions would go up. The incentive could also help fund the FTC, working to alleviate its lack of resources. This would be done by using part of the monetary penalties recovered to reward the whistleblowers and also to go either directly to the FTC, the U.S. Treasury, or all three.

The next Part confronts the potential disadvantages of a narrow incentive approach to helping the FTC. Ultimately, the next Part demonstrates that the incentive's benefits outweigh its drawbacks and that the incentive would not impede the commercial industry in a disproportionate way.

VI. DRAWBACKS OF THE INCENTIVE

While it would be convenient if a narrow FTC incentive surrounding consumer data privacy brought nothing but benefits,

165. *See* Solove & Hartzog, *supra* note 39 at 627–66.

there are potential drawbacks. Six main disadvantages are discussed below. Following each are considerations that rebut these concerns.

First, incentives used to facilitate wrongdoing can become an avenue for opportunism, over-incentivizing disclosure.[166] This carries a risk that employees might divulge conduct that is not particularly severe. Though this is a legitimate concern, overzealous claimants would still be part of a system full of checks and balances. The FTC would maintain discretion to determine whether the claims carry enough merit to bring an action. Further, the FTC would still be unable to seek monetary penalties until after a company breaches a previously entered settlement agreement. This ensures a slower pace of punishment. Also, a reward to any claimant is not a complete guarantee because there would be numerous requirements before a reward could be possible. Insiders would also still face a great amount of reputational risk outside of employer retaliation, which would deter those with imbalanced priorities.

Second, the law recognizes the importance of trade secrecy to innovation and capitalism, but this incentive could potentially play a part in destroying important proprietary information. This would happen by trade secrets becoming generally known. However, if the incentive is crafted correctly, and if insiders follow the steps to remain within its protections, insight would be given only to the FTC, meaning it would not be revealed to the media or other competitors. Further, the requirement that the insider hold a reasonable belief that the provided information relates to a possible FTC violation also helps filter out legal trade secrets from being disclosed. In sum, legal proprietary information would not easily become generally known, thereby destroying trade secrecy. Instead, only the necessary information to convey unfair or deceptive practices that are against the law would be included in publicly available information. This means that companies would still maintain legal trade secrets.

Third, the threat of employees becoming incentivized watchdogs could chill innovation. This is because companies may compartmentalize instead of collaborating, thus holding trade secrets within separate sectors to minimize potential oversight and liability. Cutting off collaboration would likely chill innovation.

166. *See* Bloch-Wehba, *supra* note 51 (manuscript at 43).

Yet this incentive is tailored much more narrowly than whistleblowing incentives already in place by the SEC, which have withstood the test of time and allowed society to continue innovating and advancing at an unprecedented pace. Additionally, incentives can be constructed to facilitate internal alerting and internal resolving before the FTC would otherwise get involved. Finally, there is a legitimate argument that not all innovation is good innovation, and it is okay for some of it to be chilled.[167] If companies sequester themselves to the point that their innovation is inhibited because their data practices are questionable, impeding this type of innovation might not be a bad result.

Fourth, some may argue the opposite problem of over-incentivizing: very few individuals actually blow the whistle to outside sources, because, as discussed above, such individuals face numerous drawbacks with high stakes.[168] However, with carefully implemented characteristics, anonymity and protections could alleviate these risks. Additionally, the incentive would not be intended to solve every problem and detect every issue for the FTC. It would only be one piece in enforcement meant to play an important but limited role in efficiently deterring wrongdoing. As previous examples demonstrate, insight from a small number of people can go a long way.

Fifth, "reactive lawmaking [alone] may . . . miss important opportunities for legal development and change."[169] However, it is important to remember that this proposed incentive is only one piece of a complex puzzle. A single incentive is not a one-size-cure-all type of solution. Nevertheless, this one piece could help regulators in the long run slowly shift from almost entirely reactive lawmaking to including some small pieces of law that are more prescriptive. A whistleblower incentive would help foster a better understanding of issues to help shape rulemaking by the FTC. This happens during the investigation process. With investigations, the FTC builds on its expertise and will begin to see firsthand what practices are prevalent, what harms are more likely, and what technologies are in play. With this new access to knowledge, regulators will be better informed to bring benefits to society.

167. *See* Stephanie Plamondon Bair, *Innovation's Hidden Externalities*, 47 BYU L. REV. 1385 (2022).

168. *See* Bloch-Wehba, *supra* note 51 (manuscript at 43).

169. *Id.* (manuscript at 45).

CONCLUSION

The status quo of the federal consumer data privacy law landscape is untenable. Society-shaping harms are at stake. Harms resulting from data privacy pitfalls are promulgated through inefficient enforcement, a lack of technological understanding, unclear legal lines, and data practices that are largely left without ethical oversight. Industry leaders and the public have called for a change, but an information imbalance between highly sophisticated companies and the lagging regulatory system has left lawmakers scratching their heads. However, at this point, an entirely prescriptive approach with such imbalanced technological understanding could overly hinder innovation by presuming how technology will evolve.

Balanced incentives are proven methods in areas of the law where trade secrecy combined with wrongdoing creates near-impossible detection. A narrow incentive with carefully crafted characteristics will help improve the status quo in numerous important ways. It would point the FTC in the right direction as it digs deep for deceptive data practices, learning and leveling the information imbalance along the way. Innovation would still flourish, and ex post enforcement would proportionately address data use to its harm. Proper data practices would become a priority, and companies would be motivated to be more transparent to limit their liability.

In sum, if left untilled and untouched, data privacy violations pose grave risks of harm to society. This Note's proposed incentive is the most realistic opportunity to impactfully address many systematic deficiencies in the current consumer data privacy landscape.